

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 11:44:08 ; Search time 39 Seconds
(without alignments)
46.875 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLNSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: Pirl:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	97	100.0	999	1 IJHUG3	desmoglein 3 precu
2	81	83.5	1043	1 IJBOG1	desmoglein 1 precu
3	81	83.5	1049	1 IJHUG1	desmoglein 1 precu
4	71	73.2	1117	2 S38673	desmoglein 2 - hum
5	52	53.6	533	2 S57904	vir49 protein - S
6	52	53.6	533	2 JC5040	positive regulator
7	52	53.6	906	1 IJHUCN	cadherin 2 precurs
8	49	50.5	877	1 IJBOCN	N-cadherin precurs
9	49	50.5	906	1 IJM5CN	N-cadherin precurs
10	48	49.5	829	2 I46536	Ksp-cadherin - rab
11	47	48.5	317	2 B96963	uncharacterized co
12	46	47.4	783	2 I50116	N-cadherin precurs
13	45	46.4	620	2 A53731	translation initia
14	45	46.4	884	1 IJM5CE	E-cadherin precurs
15	45	46.4	884	2 S34438	uvomorulin - mouse
16	45	46.4	1321	2 T23476	hypothetical prote
17	45	46.4	1321	2 S27337	multidrug resistan
18	44	45.4	347	2 D69373	immunogenic protei
19	44	45.4	512	2 T38422	probable chromatin
20	44	45.4	905	1 IJXLC1	N-cadherin 1 precu
21	44	45.4	906	1 IJXLC2	N-cadherin 2 precu
22	44	45.4	1114	2 T25083	hypothetical prote
23	44	45.4	1158	2 T25082	hypothetical prote
24	43	44.3	307	2 T40240	dimethylase - firs
25	43	44.3	315	2 T43249	rRNA (adenine-N6,N
26	43	44.3	480	2 T18675	hypothetical prote
27	43	44.3	725	1 E64211	virulence-associat
28	43	44.3	809	1 IJ80DD	desmocollin 2b pre
29	43	44.3	827	2 A53954	Li-cadherin precur

30	43	44.3	863	1 IJBODC	desmocollin 2a pre
31	43	44.3	912	1 IJCHCN	N-cadherin precurs
32	42	43.3	388	2 T09489	hypothetical prote
33	42	43.3	479	2 T30794	hypothetical prote
34	42	43.3	479	2 T28480	hypothetical prote
35	42	43.3	479	2 H72155	CLL protein - vari
36	42	43.3	479	2 E42508	ELL protein - vacc
37	42	43.3	479	2 D36841	ELL protein - vari
38	42	43.3	487	2 F84751	hypothetical prote
39	42	43.3	533	2 E86412	hypothetical prote
40	42	43.3	761	1 IJBODE	desmocollin 1a - b
41	42	43.3	770	2 B48910	desmocollin 1b pre
42	42	43.3	824	2 A48910	desmocollin 1a pre
43	42	43.3	839	1 IJBODF	desmocollin 1b pre
44	42	43.3	840	2 I37281	Decia precursor -
45	42	43.3	873	2 F86426	95.1K hypothetical

ALIGNMENTS

RESULT 1

IJHUG3

desmoglein 3 precursor - human

N:Alternate names: pemphigus vulgaris antigen

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

C:Accession: A41088

R:Amagai, M.; Klaus-Kovtun, V.; Stanley, J.R.

Cell 67, 869-877, 1991

A:Title: Autoantibodies against a novel epithelial cadherin in pemphigus vulgaris, a disease of the skin in humans

A:Reference number: A41088; PMID:1720352

A:Accession: A41088

A:Molecule type: mRNA

A:Residues: 1-999 <AMA>

A:Cross-references: UNIPROT:P32926; GB:M76482; NID:G190751; PIDN:AAA60230.1; PID:G190752

C:Genetics:

A:Gene: GDB:DSG3

A:Cross-references: GDB:134030; OMIM:169615

A:Map position: 18q12.1-18q12.2

C:Superfamily: cadherin; cadherin repeat homology

C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane protein

F:1-23/Domain: signal sequence #status predicted <Sig>

F:24-49/Domain: propeptide #status predicted <PRO>

F:50-999/Product: desmoglein homolog #status predicted <MAT>

F:50-615/Domain: extracellular #status predicted <EXT>

F:160-267/Domain: cadherin repeat homology <CR1>

F:270-383/Domain: cadherin repeat homology <CR2>

F:390-495/Domain: cadherin repeat homology <CR3>

F:496-598/Domain: cadherin repeat homology <CR4>

F:616-639/Domain: transmembrane #status predicted <TMM>

F:640-999/Domain: intracellular #status predicted <INT>

F:910-938/Domain: desmoglein repeat <DGL>

F:937-966/Domain: desmoglein repeat <DG2>

F:110,180,545/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 100.0%; Score 97; DB 1; Length 999;

Best Local Similarity 100.0%; Pred. No. 2e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLNSKIAFKIVSQEPA 19

|||||

Db 186 EPNHLNSKIAFKIVSQEPA 204

RESULT 2

IJBOG1

desmoglein 1 precursor - bovine

N:Alternate names: desmoglein BDGM

C:Species: Bos primigenius taurus (cattle)

C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

C:Accession: S14603; A38872; A37785; S38721; A48173; S24412

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S38673; B38872

R:Zimbelmann, R. EMBL Data Library, September 1993
submitted to the EMBL Data Library, September 1993

A:Reference number: S38673

A:Accession: S38673

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1117 <ZIM>

A:Cross-references: UNIPROT:Q14126; EMBL:Z26317; NID:G416177; PIDN:CAA81226.1; PID:G4161

R:Koch, P.J.; Goldschmidt, M.D.; Walsh, M.J.; Zimbelmann, R.; Franke, W.W.

Eur. J. Cell Biol. 55, 200-208, 1991

A:Title: Complete amino acid sequence of the epidermal desmoglein precursor polypeptide

A:Reference number: A38872; MUID:92037656; PMID:1935985

A:Accession: B38872

A:Molecule type: mRNA

A:Residues: 777-1117 <KOC>

A:Cross-references: GB:S64273

C:Genetics:

A:Gene: GDB.DSG2

A:Cross-references: GDB:128808; OMIM:125671

A:Map position: 18q12.1-18q12.2

C:Superfamily: cadherin; cadherin repeat homology

C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; membrane protein

F:51-158/Domain: cadherin repeat homology <CR1>

F:161-271/Domain: cadherin repeat homology <CR2>

Query Match 73.2%; Score 71; DB 2; Length 1117;
Best Local Similarity 73.7%; Pred. No. 0.00076;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEP 19

DB 187 EPTNLNLSKISYRIVSLPE 205

RESULT 5

virR49 protein - Streptococcus pyogenes (strain CS101, serotype M49)
C:Species: Streptococcus pyogenes

A:Variety: strain CS101, serotype M49

C>Date: 29-Nov-1995 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004

C:Accession: S57904; S34297; S60059

R:Podbielski, A.; Flossdorff, A.; Weber-Heynemann, J.

Infect. Immun. 63, 9-20, 1995

A:Title: The group A streptococcal virR49 gene controls expression of four structural vi

A:Reference number: S57904; MUID:95105032; PMID:7806389

A:Accession: S57904

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-533 <POD>

A:Cross-references: UNIPROT:Q54989; EMBL:X68501

R:Podbielski, A.; Weber-Heynemann, J.

submitted to the EMBL Data Library, September 1992

A:Description: VirR gene of DF-positive group A streptococci.

A:Reference number: S34297

A:Accession: S34297

A:Molecule type: DNA

A:Residues: 1-458, 'V', 460-483, 'KIKRD', 489-533 <POW>

A:Cross-references: EMBL:X68501

R:Podbielski, A.

submitted to the EMBL Data Library, February 1995

A:Reference number: S60059

A:Accession: S60059

A:Molecule type: DNA

A:Residues: 1-238, 'H', 240-329, 'D', 331-533 <POF>

A:Cross-references: EMBL:X68501; NID:G677864; PIDN:CAA48513.1; PID:G677866

C:Genetics:

A:Gene: virR49

C:Superfamily: Streptococcus positive regulatory protein Mga

Query Match 53.6%; Score 52; DB 2; Length 533;

Best Local Similarity 50.0%; Pred. No. 0.68;

Matches 10; Conservative 5; Mismatches 1; Indels 4; Gaps 1;

QY 3 NNLNSKI-----AFKIVSQEP 18

DB 269 NNLNDKLEIGCAFEINQDP 288

RESULT 6

JCS040

positive regulatory protein Mga - Streptococcus pyogenes

C:Species: Streptococcus pyogenes

C>Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 09-Jul-2004

C:Accession: JCS040; S52535

R:Andersson, G.; McIver, K.; Heden, L.O.; Scott, J.R.

Gene 175, 77-81, 1996

A:Title: Complementation of divergent mga genes in group A Streptococcus.

A:Reference number: JCS040; MUID:97074652; PMID:8917079

A:Accession: JCS040

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-533 <AND>

A:Cross-references: UNIPROT:Q54891; EMBL:X95408; NID:G1246851; PIDN:CAA64690.1; PID:G1246

A:Experimental source: strain AP4

R:Katerov, V.; Schalen, C.; Totolian, A.A.

Mol. Gen. Genet. 245, 78-85, 1994

A:Title: Sequencing of genes within the vir regulon of Streptococcus pyogenes type M15 -

A:Reference number: S52535; MUID:95147851; PMID:7845360

A:Accession: S52535

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 299-360, 'P', 362-533 <KAT>

A:Cross-references: GB:S75411; NID:G914107; PIDN:AAB33260.1; PID:G914108

C:Comment: This protein acts as the second component of a signal transducing system.

C:Genetics:

A:Gene: mga4

C:Superfamily: Streptococcus positive regulatory protein Mga

Query Match 53.6%; Score 52; DB 2; Length 533;

Best Local Similarity 50.0%; Pred. No. 0.68;

Matches 10; Conservative 5; Mismatches 1; Indels 4; Gaps 1;

QY 3 NNLNSKI-----AFKIVSQEP 18

DB 269 NNLNDKLEIGCAFEINQDP 288

RESULT 7

IJHUCN

cadherin 2 precursor - human

N:Alternate names: N-cadherin; neuronal cadherin

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

C:Accession: A38870; S11487; JQ0751; S13799

R:Reid, R.A.

submitted to the EMBL Data Library, November 1990

A:Reference number: A38870

A:Accession: A38870

A:Molecule type: mRNA

A:Residues: 1-906 <REI>

A:Cross-references: UNIPROT:P19022; EMBL:X54315; NID:G34998; PIDN:CAA38213.1; PID:G34999

Nucleic Acids Res. 18, 5896, 1990

A:Title: Human N-cadherin: nucleotide and deduced amino acid sequence.

A:Reference number: S11487; MUID:91016946; PMID:2216790

A:Accession: S11487

A:Molecule type: mRNA

A:Residues: 1-340, 'N', 342-698, 'R', 700-704, 'F', 706-906 <RE2>

A:Cross-references: EMBL:X54315

A>Note: this sequence has been revised in reference A38870

R:Walsh, F.S.; Barton, C.H.; Putt, W.; Moore, S.E.; Kelseil, D.; Spurr, N.; Goodfellow, I.

J. Neurochem. 55, 805-812, 1990

A:Title: N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin

A:Reference number: JQ0751; MUID:90347462; PMID:2384753

```
A:Accession: JQ0751
A:Molecule type: mRNA
A:Residues: 160-194,'IR',197-211,'L',213-227,'Q',229,'N',231-235,'G',237-248,'T',250-356
A:Cross-references: GB:M34064
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to
C:Genetics:
A:Gene: GDB:CDH2; NCAD
A:Cross-references: GDB:128185; OMIM:114020
A:Map position: 18q12.1-18q12.1
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pr
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-159/Domain: propeptide #status predicted <PRO>
F:160-906/Product: N-cadherin #status predicted <MAT>
F:160-714/Domain: extracellular #status predicted <EXT>
F:162-267/Domain: cadherin repeat homology <CR1>
F:237-242/Region: cadherin binding #status predicted
F:270-382/Domain: cadherin repeat homology <CR2>
F:388-497/Domain: cadherin repeat homology <CR3>
F:500-605/Domain: cadherin repeat homology <CR4>
F:606-712/Domain: cadherin repeat homology <CR5>
F:715-746/Domain: transmembrane #status predicted <TMM>
F:747-906/Domain: intracellular #status predicted <INT>
F:865-878/Region: serine-rich
F:190,273,325,402,572,651,692/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 53.6%; Score 52; DB 1; Length 906;
Best Local Similarity 47.4%; Pred. No. 1.2;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHNSKIAPKIVSQEPA 19
Db 296 DPNALNGMLRYRILSQAPS 314
: || | : : || | :

RESULT 8
IJBON
N-cadherin precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: S11693
R:Liew, C.W.; Cannon, C.; Power, M.D.; Kiboneka, P.K.; Rubin, L.L.
EMBO J. 9, 2701-2708, 1990
A:Title: Identification and cloning of two species of cadherins in bovine endothelial ce
A:Reference number: S11693; MUID:90360979; PMID:2390969
A:Accession: S11693
A:Molecule type: mRNA
A:Residues: 1-877 <LIR>
A:Cross-references: UNIPROT:P19534; EMBL:X53615; NID:g164; PIDN:CAA37677.1; PID:g664894
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought b
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pr
F:1-130/Domain: propeptide (fragment) #status predicted <PRO>
F:131-877/Product: N-cadherin #status predicted <MAT>
F:131-685/Domain: extracellular #status predicted <EXT>
F:133-238/Domain: cadherin repeat homology <CR1>
F:208-213/Region: cadherin binding #status predicted
F:241-353/Domain: cadherin repeat homology <CR2>
F:356-468/Domain: cadherin repeat homology <CR3>
F:471-576/Domain: cadherin repeat homology <CR4>
F:577-685/Domain: cadherin repeat homology <CR5>
F:686-717/Domain: transmembrane #status predicted <TMM>
F:718-877/Domain: intracellular #status predicted <INT>
F:836-849/Region: serine-rich
F:161,244,296,373,543,593,622,663/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 50.5%; Score 49; DB 1; Length 877;
Best Local Similarity 42.1%; Pred. No. 3.9;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHNSKIAPKIVSQEPA 19
Db 267 DPNALNGMLRYRILSQAPS 285
: || | : : || | :

A:Accession: JQ0751
A:Molecule type: mRNA
A:Residues: 160-194,'IR',197-211,'L',213-227,'Q',229,'N',231-235,'G',237-248,'T',250-356
A:Cross-references: GB:M34064
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to
C:Genetics:
A:Gene: GDB:CDH2; NCAD
A:Cross-references: GDB:128185; OMIM:114020
A:Map position: 18q12.1-18q12.1
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pr
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-159/Domain: propeptide #status predicted <PRO>
F:160-906/Product: N-cadherin #status predicted <MAT>
F:160-714/Domain: extracellular #status predicted <EXT>
F:162-267/Domain: cadherin repeat homology <CR1>
F:237-242/Region: cadherin binding #status predicted
F:270-382/Domain: cadherin repeat homology <CR2>
F:388-497/Domain: cadherin repeat homology <CR3>
F:500-605/Domain: cadherin repeat homology <CR4>
F:606-712/Domain: cadherin repeat homology <CR5>
F:715-746/Domain: transmembrane #status predicted <TMM>
F:747-906/Domain: intracellular #status predicted <INT>
F:865-878/Region: serine-rich
F:190,273,325,402,572,651,692/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 53.6%; Score 52; DB 1; Length 906;
Best Local Similarity 47.4%; Pred. No. 1.2;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHNSKIAPKIVSQEPA 19
Db 296 DPNALNGMLRYRILSQAPS 314
: || | : : || | :

RESULT 8
IJBON
N-cadherin precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: S11693
R:Liew, C.W.; Cannon, C.; Power, M.D.; Kiboneka, P.K.; Rubin, L.L.
EMBO J. 9, 2701-2708, 1990
A:Title: Identification and cloning of two species of cadherins in bovine endothelial ce
A:Reference number: S11693; MUID:90360979; PMID:2390969
A:Accession: S11693
A:Molecule type: mRNA
A:Residues: 1-877 <LIR>
A:Cross-references: UNIPROT:P19534; EMBL:X53615; NID:g164; PIDN:CAA37677.1; PID:g664894
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought b
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pr
F:1-130/Domain: propeptide (fragment) #status predicted <PRO>
F:131-877/Product: N-cadherin #status predicted <MAT>
F:131-685/Domain: extracellular #status predicted <EXT>
F:133-238/Domain: cadherin repeat homology <CR1>
F:208-213/Region: cadherin binding #status predicted
F:241-353/Domain: cadherin repeat homology <CR2>
F:356-468/Domain: cadherin repeat homology <CR3>
F:471-576/Domain: cadherin repeat homology <CR4>
F:577-685/Domain: cadherin repeat homology <CR5>
F:686-717/Domain: transmembrane #status predicted <TMM>
F:718-877/Domain: intracellular #status predicted <INT>
F:836-849/Region: serine-rich
F:161,244,296,373,543,593,622,663/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 50.5%; Score 49; DB 1; Length 877;
Best Local Similarity 42.1%; Pred. No. 3.9;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHNSKIAPKIVSQEPA 19
Db 267 DPNALNGMLRYRILSQAPS 285
: || | : : || | :
```

RESULT 9

IJMSCN

N-cadherin precursor, neuronal - mouse
C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: A32759; A46163

R:Miyatani, S.; Shimamura, K.; Hatta, M.; Nagafuchi, A.; Nose, A.; Matsunaga, M.; Hatta, M.; Science 245, 631-635, 1989

A:Title: Neural cadherin: role in selective cell-cell adhesion.

A:Reference number: A32759; MUID:83346748; PMID:2762814

A:Accession: A32759

A:Molecule type: mRNA

A:Residues: 1-906 <MIY>

A:Cross-references: UNIPROT:P15116; GB:M31131; NID:g192327; PIDN:AAA37353.1; PID:g309125
R:Miyatani, S.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Takeichi, M.

Proc. Natl. Acad. Sci. U.S.A. 89, 8443-8447, 1992

A:Title: Genomic structure and chromosomal mapping of the mouse N-cadherin gene.

A:Reference number: A46163; MUID:92409532; PMID:1528849

A:Accession: A46163

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 839-906 <MI2>

A:Cross-references: GB:S45011; NID:g256010; PIDN:AB23356.1; PID:g256011

A>Note: sequence extracted from NCBI backbone (NCBIN:113759, NCBIP:113760)

C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Superfamily: cadherin; cadherin repeat homology

C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane prot
F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-159/Domain: propeptide #status predicted <PRO>

F:160-906/Product: N-cadherin #status predicted <MAT>

F:160-714/Domain: extracellular #status predicted <EXT>

F:162-267/Domain: cadherin repeat homology <CR1>

F:237-242/Region: cadherin binding #status predicted

F:270-382/Domain: cadherin repeat homology <CR2>

F:388-497/Domain: cadherin repeat homology <CR3>

F:500-605/Domain: cadherin repeat homology <CR4>

F:606-714/Domain: cadherin repeat homology <CR5>

F:715-746/Domain: transmembrane #status predicted <TMM>

F:747-906/Domain: intracellular #status predicted <INT>

F:865-878/Region: serine-rich

F:190,273,325,402,572,651,692/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match Similarity 50.5%; Score 49; DB 1; Length 906;

Best Local Similarity 42.1%; Pred. No. 4.1;

Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHNSKIAPKIVSQEPA 19

Db 296 DPNALNGMLRYRILSQAPS 314

: || | : : || | :

RESULT 10

Ksp-cadherin - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004

C:Accession: I46536

R:Thomson, R.B.; Igarashi, P.; Bienesderfer, D.; Kim, R.; Abu-Alfa, A.; Soleimani, M.; A

J. Biol. Chem. 270, 17594-17601, 1995

A:Title: Isolation and cDNA cloning of Ksp-cadherin, a novel kidney-specific member of t

A:Reference number: I46536; MUID:95340560; PMID:7615566

A:Accession: I46536

A>Status: preliminary;

A:Molecule type: mRNA

A:Residues: 1-829 <THO>

A:Cross-references: UNIPROT:Q28634; EMBL:U28945; NID:g902885; PIDN:AAC48472.1; PID:g90288

Query Match 49.5%; Score 48; DB 2; Length 829;

Best Local Similarity 52.6%; Pred. No. 5.5;

Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
||| : : : : :
Db 155 EPGTANSURFILLSQTPA 173

RESULT 11
B96963
uncharacterized conserved protein, similar to B. subtilis yvCL CAC0513 [imported] - Clo
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: B96963
R:Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B96963
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-317 <KUR>
A:Cross-references: UNIPROT:Q97LP1; GB:AE001437; PID:AAK78493.1; PID:g15023376; GSPDB:C
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
C:Superfamily: Streptomyces coelicolor hypothetical protein SCC54.10c

Query Match 48.5%; Score 47; DB 2; Length 317;
Best Local Similarity 57.1%; Pred. No. 2.8;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQEPA 19
| : : : : :
Db 41 NRKISFKIITENPA 54

RESULT 12
I50116
N-cadherin precursor - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jan-2000
C:Accession: I50116
R:Bitzur, S.; Kam, Z.; Geiger, B.
Dev. Dyn. 201, 121-136, 1994
A:Title: Structure and distribution of N-cadherin in developing zebrafish embryos: morph
A:Reference number: I50116; MUID:95178741; PMID:7873785
A:Accession: I50116
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-783 <BIT>
A:Cross-references: EMBL:X67648; NID:G623209; PID:CAA47890.1; PID:g623210
C:Superfamily: cadherin; cadherin repeat homology
F:146-258/Domain: cadherin repeat homology <CDH>

Query Match 47.4%; Score 46; DB 2; Length 783;
Best Local Similarity 44.4%; Pred. No. 11;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEP 18
: : : : :
Db 172 DENTANGMLRYKILLSQTP 189

RESULT 13
A53731
translation initiation factor eIF-2 alpha chain kinase (EC 2.7.1.-) - rat
N:Alternate names: heme-controlled repressor
C:Species: Rattus norvegicus (Norway rat)
C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C:Accession: A53731
R:Mellor, H.; Flowers, K.M.; Kimball, S.R.; Jefferson, L.S.
J. Biol. Chem. 269, 10201-10204, 1994
A:Title: Cloning and characterization of cDNA encoding rat hemin-sensitive initiation fa
A:Reference number: A53731; MUID:94193700; PMID:7908290

A:Accession: A53731
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-620 <MEL>
A:Cross-references: UNIPROT:Q63185; GB:L27707; NID:G443688; PIDN:AAA18255.1; PID:g443689
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: phosphotransferase
F:165-581/Domain: protein kinase homology <KIN>

Query Match 46.4%; Score 45; DB 2; Length 620;
Best Local Similarity 56.2%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQ 16
||| : : : : :
Db 81 EPNPLHSKQVFKLLCQ 96

RESULT 14
IUMSCE
E-cadherin precursor, epithelial - mouse
N:Alternate names: uvomorulin
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: S04528; S03160; I49565; S48735
R:Nagafuchi, A.; Shiryayoshi, Y.; Okazaki, K.; Yasuda, K.; Takeichi, M.
Nature 329, 341-343, 1987
A:Title: Transformation of cell adhesion properties by exogenously introduced E-cadherin
A:Reference number: S04528; MUID:87315445; PMID:3498123
A:Accession: S04528
A:Molecule type: mRNA
A:Residues: 1-412, V, 414-884 <NAG>
A:Cross-references: UNIPROT:P09803; EMBL:X06115
R:Ringwald, M.; Schuh, R.; Vestweber, D.; Eistetter, H.; Lottspeich, F.; Engel, J.; Doeli,
EMBO J. 6, 3647-3653, 1987
A:Title: The structure of cell adhesion molecule uvomorulin. Insights into the molecular
A:Reference number: S03160; MUID:8811553; PMID:3501370
A:Accession: S03160
A:Molecule type: mRNA
A:Residues: 157-884 <RIN>
A:Cross-references: EMBL:X06339
A:Note: part of this sequence, including the amino end of the mature protein, was confirm
R:Behrens, J.; Lowrick, O.; Klein-Hitpass, L.; Birchmeier, W.
Proc. Natl. Acad. Sci. U.S.A. 88, 11495-11499, 1991
A:Title: The E-cadherin promoter: Functional analysis of a G-C-rich region and an epithe
A:Reference number: I49565; MUID:92107977; PMID:1763063
A:Accession: I49565
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-15 <RES>
A:Cross-references: GB:M81449; NID:g192325; PIDN:AAA37352.1; PID:g192326
R:Tong, K.I.; Yau, P.; Overduin, M.; Bagby, S.; Porumb, T.; Takeichi, M.; Ikura, M.
PDBS Lett. 352, 318-322, 1994
A:Title: Purification and spectroscopic characterization of a recombinant amino-terminal
A:Reference number: S48735; MUID:95010732; PMID:7925993
A:Accession: S48735
A:Status: preliminary
A:Molecule type: protein
A:Residues: 156-300 <TON>
A:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Genetics:
A:Gene: E-cadherin
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane prot
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-156/Domain: propeptide #status predicted <PRO>
F:157-884/Product: E-cadherin, epithelial #status experimental <MAT>
F:157-699/Domain: extracellular #status predicted <EXT>
F:159-264/Domain: cadherin repeat homology <CR1>
F:234-239/Region: cadherin binding #status predicted
F:267-377/Domain: cadherin repeat homology <CR2>
F:380-486/Domain: cadherin repeat homology <CR3>
F:489-597/Domain: cadherin repeat homology <CR4>

F;598-702/Domain: cadherin repeat homology <CR5>
F;702-733/Domain: transmembrane #status predicted <TM>
F;734-884/Domain: intracellular #status predicted <INT>
F;842-855/Region: serine-rich
F;560,639/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 46.4%; Score 45; DB 1; Length 884;
Best Local Similarity 56.2%; Pred. No. 20;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQEP 18
| | | | | | | | | |
DB 296 NTYNAAIAYTVSQDP 311

RESULT 15
S34438
uvomorulin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S34438
R;Ringwald, M.; Baribault, H.; Schmidt, C.; Kemler, R.
Nucleic Acids Res. 19, 6533-6539, 1991
A;Title: The structure of the gene coding for the mouse cell adhesion molecule uvomorulin
A;Reference number: S34438; MUID:92093614; PMID:1754391
A;Accession: S34438
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-884 <RIN>
A;Cross-references: UNIPROT:P09803; EMBL:X60975
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
C;Genetics:
A;Introns: 16/3; 57/1; 131/3; 179/3; 231/3; 280/1; 338/3; 381/3; 442/3; 524/2; 573/1; 64
C;Superfamily: cadherin; cadherin repeat homology
F;380-488/Domain: cadherin repeat homology <CR3>

Query Match 46.4%; Score 45; DB 2; Length 884;
Best Local Similarity 56.2%; Pred. No. 20;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQEP 18
| | | | | | | | | |
DB 296 NTYNAAIAYTVSQDP 311

Search completed: March 15, 2005, 11:54:38
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 11:32:37 ; Search time 174 Seconds
(without alignments)
55.917 Million cell updates/sec

Title: US-10-799-005A-1
Perfect score: 97
Sequence: 1 EPNHLSKIAPKIVSQEPA 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	97	100.0	993	2	Q7YRU7	Q7YRU7 canis famil
2	97	100.0	999	1	DSG3_HUMAN	P32926 homo sapien
3	95	97.9	929	2	Q8CE48	Q8CE48 mus musculu
4	95	97.9	993	1	DSG3_MOUSE	O35302 mus musculu
5	95	97.9	993	2	Q8CB02	Q8CB02 mus musculu
6	85	87.6	1040	2	Q6W3B0	Q6W3B0 rattus norv
7	85	87.6	1041	2	Q7TMD7	Q7TMD7 mus musculu
8	83	85.6	1040	2	Q86S06	Q86S06 mus musculu
9	83	85.6	1059	2	Q6Y9L9	Q6Y9L9 homo sapien
10	81	83.5	1043	1	DSG1_BOVIN	Q03763 bos taurus
11	81	83.5	1049	1	DSG1_HUMAN	Q02413 homo sapien
12	81	83.5	1054	2	Q9GKQ8	Q9GKQ8 canis famil
13	75	77.3	911	2	Q7TQ61	Q7TQ61 mus musculu
14	75	77.3	911	2	Q7TQF0	Q7TQF0 mus musculu
15	75	77.3	1060	2	Q7TQ60	Q7TQ60 mus musculu
16	75	77.3	1060	2	Q7TQF1	Q7TQF1 mus musculu
17	71	73.2	1117	1	DSG2_HUMAN	P19022 homo sapien
18	65	67.0	392	2	Q8VCE3	Q8VCE3 mus musculu
19	65	67.0	360	2	Q8L111	Q8L111 mus musculu
20	65	67.0	1122	1	DSG2_MOUSE	O55111 mus musculu
21	54	55.7	639	2	Q7YQF5	Q7YQF5 oryza sativ
22	54	55.7	864	2	Q7XVV5	Q7XVV5 oryza sativ
23	53	54.6	200	2	Q8SD26	Q8SD26 pseudononas
24	52	53.6	533	2	Q54891	Q54891 streptococc
25	52	53.6	533	2	Q54989	Q54989 streptococc
26	52	53.6	533	2	Q8NZ78	Q8NZ78 streptococc
27	52	53.6	906	1	CAD2_HUMAN	P19022 homo sapien
28	52	53.6	906	1	Q8N173	Q8N173 homo sapien
29	50	51.5	742	2	Q6GL53	Q6GL53 xenopus tro
30	49	50.5	238	1	CAD2_CRIGR	O55075 cricetus
31	49	50.5	707	2	Q985J8	Q985J8 rhizobium 1

32	49	50.5	877	1	CAD2_BOVIN	P19534 bos taurus
33	49	50.5	906	1	CAD2_MOUSE	P15116 mus musculu
34	49	50.5	906	1	CAD2_RAT	Q921y3 rattus norv
35	49	50.5	906	2	Q8BS19	Q8BS19 mus musculu
36	48	49.5	807	2	Q6UW93	Q6UW93 homo sapien
37	48	49.5	829	1	CADG_HUMAN	O75309 homo sapien
38	48	49.5	829	1	CADG_RABIT	Q28634 oryctolagus
39	47.5	49.0	464	2	Q89QR0	Q89QR0 bradyrhizob
40	47	48.5	317	2	Q97LPI	Q97LPI clostridium
41	46	47.4	249	2	Q6DC16	Q6DC16 brachydanio
42	46	47.4	299	1	HIS1_PASMU	P57919 pasteurella
43	46	47.4	334	2	Q9Z643	Q9Z643 enterobacte
44	46	47.4	827	1	CADH_MOUSE	Q9R100 mus musculu
45	46	47.4	893	1	CAD2_BRARE	Q90275 brachydanio

ALIGNMENTS

RESULT 1

Q7YRU7 PRELIMINARY; PRT; 993 AA.
AC Q7YRU7;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Desmoglein 3.
GN Name=DSG3;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBITaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oral mucosa;
RA Nishifuji K., Amagai M., Ota T., Park S.-J., Nishikawa T., Iwasaki T.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins
(By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Contains 2 cadherin domains.
DR EMBL; AF394784; RAP80592.1; -.
DR HSSP; P09803; 1FF5
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00028; Cadherin; 4.
DR Pfam; PF01049; Cadherin_C; 1.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 2.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN 1.
KW Calcium; Calcium-binding; Cell adhesion; Transmembrane.
SQ SEQUENCE 993 AA; 107550 MW; BIEE9C6D59666D5D CRC64;

Query Match 100.0%; Score 97; DB 2; Length 993;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
Db 185 EPNHLSKIAPKIVSQEPA 203
|||||

RESULT 2

DSG3_HUMAN STANDARD; PRT; 999 AA.
ID_DSG3_HUMAN
AC P32926;

01-OCT-1993 (Rel. 27, Created)
01-OCT-1993 (Rel. 27, Last sequence update)
05-JUL-2004 (Rel. 44, Last annotation update)
Desmoglein 3 precursor (130 kDa pemphigus vulgaris antigen) (PVA).
GN Name=DSG3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92069753; PubMed=1720352; DOI=10.1016/0092-8674(91)90360-B;
RA Amagai M., Klaus-Kovtun V., Stanley J.R.;
RT "Autoantibodies against a novel epithelial cadherin in pemphigus
vulgaris, a disease of cell adhesion.";
RL Cell 67:869-877(1991).
CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved
CC in the interaction of plaque proteins and intermediate filaments
CC mediating cell-cell adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Epidermis, tongue, tonsil, esophagus and
CC carcinomas.
CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
CC (potential).
CC -!- DISEASE: Pemphigus vulgaris (PV) is a potentially lethal skin
CC disease in which epidermal blisters occur as the result of the
CC loss of cell-cell adhesion caused by the action of autoantibodies
CC against desmoglein 3.
CC -!- SIMILARITY: Contains 4 cadherin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M76482; AAA60230.1; --
DR PIR; A41088; IJHUG3.
DR HSP; P15116; INCG3.
DR Genew; HGNC:3050; DSG3.
DR MIM; 169615; --
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR009122; Desmo cadherin.
DR InterPro; IPR009123; Desmoglein.
DR Pfam; PF00028; Cadherin.4.
DR PRINTS; PR00205; CADHERIN.
DR PRINTS; PR01818; DESMOCADHERN.
DR PRINTS; PR01819; DESMOGLEIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 4.
KW Calcium-binding; Cell adhesion; Cytoskeleton; Glycoprotein; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 23
FT PROPEP 24 49 Potential.
FT CHAIN 50 999 Desmoglein 3.
FT DOMAIN 50 615 Extracellular (Potential).
FT TRANSMEM 616 640 Potential.
FT DOMAIN 641 999 Cytoplasmic (Potential).
FT DOMAIN 50 158 Cadherin 1.
FT DOMAIN 159 268 Cadherin 2.
FT DOMAIN 269 383 Cadherin 3.
FT DOMAIN 386 499 Cadherin 4.
FT REPEAT 910 935 Desmoglein repeat 1.
FT REPEAT 936 966 Desmoglein repeat 2.
FT CARBOHYD 110 110 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 180 180 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 459 459 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 545 545 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 999 999 AA; 107503 MW; 60479DD46AC219A1 CRC64;
SEQUENCE

Query Match 100.0%; Score 97; DB 1; Length 999;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPNHLSKIAPKIYSQEP 19
DB 186 EPNHLSKIAPKIYSQEP 204
RESULT 3
QSC48 PRELIMINARY; PRT; 929 AA.
AC Q8CE48;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
DE library, clone:4732482D04 product:desmoglein 3, full insert
DE sequence.
GN Name=DSG3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,


```

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saichoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Tota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 2 cadherin domains.
DR EMBL; AK029018; BAC36245.1; -.
DR HSP; P15116; INCUJ.
DR MGD; MGI:99499; Dsg3.
DR GO; GO:0030057; C:desmosome; TAS.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005914; C:spot adherens junction; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR009123; Desmoglein.
DR InterPro; IPR009122; Desmo.cadherin.
DR Pfam; PF00028; Cadherin; 4.
DR PRINTS; PR00205; CADHERIN.
DR PRINTS; PR01818; DESMOCADHERN.
DR PRINTS; PR01819; DESMOGLEIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS02268; CADHERIN_2; 4.
KW Calcium-binding; Cell adhesion; Cytoskeleton; Glycoprotein; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 23 Potential.
FT PROPEP 24 49 Potential.
FT CHAIN 50 >993 Desmoglein 3. (Potential).
FT DOMAIN 24 617 Extracellular (Potential).
FT TRANSMEM 618 638 Potential.
FT DOMAIN 639 993 Cytoplasmic (Potential).
FT DOMAIN 49 157 Cadherin 1.
FT DOMAIN 158 267 Cadherin 2.
FT DOMAIN 268 388 Cadherin 3.
FT DOMAIN 384 495 Cadherin 4.
FT REPEAT 905 930 Desmoglein repeat 1.
FT REPEAT 931 961 Desmoglein repeat 2.
FT CARBOHYD 110 110 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 180 180 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 459 459 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 546 546 N-linked (GlcNAc . .) (Potential).
FT NON_TER 993 993
SQ SEQUENCE 929 AA; 101357 MW; 6B38BE7B68454BA6 CRC64;

Query Match 97.9%; Score 95; DB 2; Length 929;
Best Local Similarity 94.7%; Pred. No. 3.3e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHNSKIAPKIVSQEPA 19
Db ||||:|||||
186 EPNHNSKIAPKIVSQEPA 204

RESULT 4
DSG3_MOUSE STANDARD; PRT; 993 AA.
AC O35902;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Desmoglein 3 precursor (130 kDa pemphigus vulgaris antigen homolog)
DE (Fragment).
GN Name=Dsg3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95201440; PubMed=7894164;
RA Ishikawa H., Silos S.A., Tamai K., Copeland N.G., Gilbert D.J.,
RT "cDNA cloning and chromosomal assignment of the mouse gene for
RT desmoglein 3 (Dsg3), the pemphigus vulgaris antigen."
RL Mamm. Genome 5:803-804(1994).
CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved
CC in the interaction of plaque proteins and intermediate filaments
CC mediating cell-cell adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
CC (Potential).
CC -!- SIMILARITY: Contains 4 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U06016; A065091.1; -.
DR HSP; P15116; INCUJ.
DR MGD; MGI:99499; Dsg3.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR009122; Desmo.cadherin.
DR InterPro; IPR009123; Desmoglein.
DR Pfam; PF00028; Cadherin; 4.
DR PRINTS; PR00205; CADHERIN.
DR PRINTS; PR01818; DESMOCADHERN.
DR PRINTS; PR01819; DESMOGLEIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS02268; CADHERIN_2; 4.
KW Calcium-binding; Cell adhesion; Cytoskeleton; Glycoprotein; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 23 Potential.
FT PROPEP 24 49 Potential.
FT CHAIN 50 >993 Desmoglein 3. (Potential).
FT DOMAIN 24 617 Extracellular (Potential).
FT TRANSMEM 618 638 Potential.
FT DOMAIN 639 993 Cytoplasmic (Potential).
FT DOMAIN 49 157 Cadherin 1.
FT DOMAIN 158 267 Cadherin 2.
FT DOMAIN 268 388 Cadherin 3.
FT DOMAIN 384 495 Cadherin 4.
FT REPEAT 905 930 Desmoglein repeat 1.
FT REPEAT 931 961 Desmoglein repeat 2.
FT CARBOHYD 110 110 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 180 180 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 459 459 N-linked (GlcNAc . .) (Potential).
FT CARBOHYD 546 546 N-linked (GlcNAc . .) (Potential).
FT NON_TER 993 993
SQ SEQUENCE 993 AA; 107888 MW; DB6CC526ABFB179A CRC64;

Query Match 97.9%; Score 95; DB 1; Length 993;
Best Local Similarity 94.7%; Pred. No. 3.6e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHNSKIAPKIVSQEPA 19
Db ||||:|||||
186 EPNHNSKIAPKIVSQEPA 204

RESULT 5
Q8CB02 PRELIMINARY; PRT; 993 AA.
AC Q8CB02;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Mus musculus adult female vagina cDNA, RIKEN full-length enriched
DE library, clone:9930125N18 product:desmoglein 3, full insert
DE sequence.
GN Name=Dsg3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RT Meth. Enzymol. 303:19-44(1999).
RN [2]

```



```
RESULT 7
Q7TMD7
ID Q7TMD7 PRELIMINARY; PRT; 1041 AA.
AC Q7TMD7;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Desmoglein 4.
GN Name=DSG4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22672408;
RA Whitlock N.V.;
RT "Genetic analysis of the mouse desmoglein cluster reveals evidence for
RT six distinct genes: Characterization of Mouse DSG4, DSG5, and DSG6.";
RL J. Invest. Dermatol. 120:970-980(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PMK;
RX MEDLINE=22592666; PubMed=12705872; DOI=10.1016/S0092-8674(03)00273-3;
RA Kljuic A., Bazzi H., Sundberg J.P., Martinez-Mir A., O'Shaughnessy R.,
RA Mahoney M.G., Levy M., Montagutelli X., Ahmad W., Alta V.M.,
RA Gordon D., Uitto J., Whiting D., Ott J., Fischer S., Gilliam T.C.,
RA Jahoda C.A., Morris R.J., Panteleyev A.A., Nguyen V.T.,
RA Cristiano A.M.;
RT "Desmoglein 4 in hair follicle differentiation and epidermal adhesion.
RT Evidence from inherited hypotrichosis and acquired pemphigus
RT vulgaris.";
RL Cell 113:249-260(2003).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Contains 4 cadherin domains.
DR EMBL; AY191584; AAO91793.1; -.
DR EMBL; AY227349; AAP44999.1; -.
DR HSP; P09803; 1FF5.
DR GO; GO:0030057; C:desmosome; IMP.
DR GO; GO:0042640; P:anagen; IMP.
DR GO; GO:0016337; P:cell-cell adhesion; IMP.
DR GO; GO:0030216; P:keratinocyte differentiation; IMP.
DR GO; GO:0002028; Cadherin; 4.
DR InterPro; IPR002126; Cadherin_C_term.
DR Pfam; PF01049; Cadherin_C; 1.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 4.
KW Calcium; Calcium-binding; Cell adhesion; Transmembrane.
SQ SEQUENCE 1041 AA; 114448 MW; 230ECC0553048B2B CRC64;

Query Match 87.6%; Score 85; DB 2; Length 1041;
Best Local Similarity 84.2%; Pred. No. 28-05;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
DB 186 EDNHLNSKIAYKIVSQEPA 204

RESULT 8
Q86SJ6
ID Q86SJ6 PRELIMINARY; PRT; 1040 AA.
AC Q86SJ6;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Desmoglein 4 preproprotein precursor.
GN Name=DSG4;
OS Homo sapiens (Human).

RESULT 9
QY9L9
ID QY9L9 PRELIMINARY; PRT; 1059 AA.
AC QY9L9;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Desmoglein 4.
GN Name=DSG4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22536008; PubMed=12648213;
RA Whitlock N.V., Bower C.;
RT "Genetic evidence for a novel human desmosomal cadherin, desmoglein
RT 4.";
RL J. Invest. Dermatol. 120:523-530(2003).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22592666; PubMed=12705872; DOI=10.1016/S0092-8674(03)00273-3;
RA Kljuic A., Bazzi H., Sundberg J.P., Martinez-Mir A., O'Shaughnessy R.,
RA Mahoney M.G., Levy M., Montagutelli X., Ahmad W., Alta V.M.,
RA Gordon D., Uitto J., Whiting D., Ott J., Fischer S., Gilliam T.C.,
RA Jahoda C.A., Morris R.J., Panteleyev A.A., Nguyen V.T.,
RA Cristiano A.M.;
RT "Desmoglein 4 in hair follicle differentiation and epidermal adhesion.
RT Evidence from inherited hypotrichosis and acquired pemphigus
RT vulgaris.";
RL Cell 113:249-260(2003).
CC -!- SIMILARITY: Contains 4 cadherin domains.
DR EMBL; AY177664; AAO47078.1; -.
DR EMBL; AY227350; AAP45000.1; -.
DR HSP; AY177663; AAO47077.1; -.
DR HSP; P15116; INCU.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR009123; Desmoglein.
DR InterPro; IPR009122; Desmo_cadherin.
DR Pfam; PF00028; Cadherin; 4.
DR PRINTS; PR00205; CADHERIN.
DR PRINTS; PR01818; DESMOCADHERN.
DR PRINTS; PR01819; DESMOGLEIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 4.
KW Calcium; Calcium-binding; Signal.
FT SIGNAL 1 21 Potential.
FT CHAIN 50 1040 desmoglein 4.
SQ SEQUENCE 1040 AA; 113823 MW; 0385401584C55CF6 CRC64;

Query Match 85.6%; Score 83; DB 2; Length 1040;
Best Local Similarity 84.2%; Pred. No. 4.3e-05;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
DB 186 EDNHLNSKIAYKIVSQEPA 204
```

RL ANIMAL MODELS OF HUMAN INFLAMMATORY SKIN DISEASES, pp.0-0, CRC Press
RL LLC, FL, USA (2003).
RP [2]
RA Nguyen V.T.;
RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 4 cadherin domains.
DR EMBL; AY168788; AA043657.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR009123; Desmoglein.
DR InterPro; IPR009122; Desmo_cadherin.
DR Pfam; PF00028; Cadherin; 4.
DR PRINTS; PR00205; CADHERIN.
DR PRINTS; PR01818; DESMOCADHERN.
DR PRINTS; PR01819; DESMOGLEIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 4.
KW Calcium; Calcium-binding.
SQ SEQUENCE 1059 AA; 115450 MW; EA475FDA4C2244624 CRC64;

Query Match 85.6%; Score 83; DB 2; Length 1059;
Best Local Similarity 84.2%; Pred. No. 4.4e-05;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHNSKIATKIVSQEPA 19
DB 186 EENHLNSKIATKIVSQEPS 204

RESULT 10
DSGL_BOVIN STANDARD; PRT; 1043 AA.
AC Q03763; Q9T515;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Desmoglein 1 precursor (Desmosomal glycoprotein 1) (DGI) (DGI)
DE (Pemphigus foliaceus antigen).
GN Name=DSGL;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muzzle epithelium;
RA Koch P.J., Goldschmidt M.D., Zimbelmann R., Franke W.W.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RA Puttagunta S., Mathur M., Cowin P.;
RL "Structure of DSGL, the bovine desmosomal cadherin gene encoding the pemphigus foliaceus antigen. Evidence of polymorphism.";
RL J. Biol. Chem. 269:1949-1955 (1994).
RN [3]
RP SEQUENCE OF 44-1043 FROM N.A.
RC TISSUE=Muzzle epithelium;
RX MEDLINE=91168965; PubMed=1706270;
RA Koch P.J., Walsh M.J., Schmelz M., Goldschmidt M.D., Zimbelmann R.,
RA Franke W.W.;
RT "Identification of desmoglein, a constitutive desmosomal glycoprotein, as a member of the cadherin family of cell adhesion molecules.";
RL Eur. J. Cell Biol. 53:1-12 (1990).
RN [4]
RP SEQUENCE OF 101-123, AND REVISIONS.
RX MEDLINE=92037656; PubMed=1935985;

RA Koch P.J., Goldschmidt M.D., Walsh M.J., Zimbelmann R., Franke W.W.;
RT "Complete amino acid sequence of the epidermal desmoglein precursor polypeptide and identification of a second type of desmoglein gene.";
RL Eur. J. Cell Biol. 55:200-208 (1991).
RN [5]
RP SEQUENCE OF 44-493 FROM N.A.
RX MEDLINE=91097553; PubMed=1702628;
RA Goodwin L., Hill J.E., Raynor K., Raszi L., Manabe M., Cowin P.;
RT "Desmoglein shows extensive homology to the cadherin family of cell adhesion molecules.";
RL Biochem Biophys. Res. Commun. 173:1224-1230 (1990).
CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved in the interaction of plaque proteins and intermediate filaments mediating cell-cell adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Epidermis, muzzle, tongue and esophagus.
CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats (Potential).
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58466; CAA41380.1; -;
DR EMBL; X57784; CAA40930.1; -;
DR EMBL; M58165; AAG62709.1; -;
DR PIR; S14603; IJBOGI.
DR HSSP; P15116; INCUJ.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin C term.
DR InterPro; IPR009122; Desmo_cadherin.
DR InterPro; IPR009123; Desmoglein.
DR Pfam; PF00028; Cadherin; 3.
DR Pfam; PF01049; Cadherin C; 1.
DR PRINTS; PR00205; CADHERIN.
DR PRINTS; PR01818; DESMOCADHERN.
DR PRINTS; PR01819; DESMOGLEIN.
DR SMART; SM00112; CA; 3.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 4. Cytoskeleton;
KW Calcium-binding; Cell adhesion; Cytoskeleton;
KW Direct protein sequencing; Glycoprotein; Repeat; Signal;
KW Transmembrane.
FT SIGNAL 1 23 Potential.
FT PROPEP 24 49 Potential.
FT CHAIN 50 1043 Desmoglein 1.
FT DOMAIN 50 548 Extracellular (Potential).
FT TRANSMEM 549 573 Potential.
FT DOMAIN 574 1043 Cytoplasmic (Potential).
FT DOMAIN 50 158 Cadherin 1.
FT DOMAIN 159 270 Cadherin 2.
FT DOMAIN 271 385 Cadherin 3.
FT DOMAIN 386 498 Cadherin 4.
FT DOMAIN 499 845 Desmoglein repeat 1.
FT REPEAT 819 845 Desmoglein repeat 2.
FT REPEAT 846 875 Desmoglein repeat 3.
FT REPEAT 876 905 Desmoglein repeat 4.
FT REPEAT 906 933 Desmoglein repeat 5.
FT REPEAT 934 962 Gly/Ser-rich.
FT DOMAIN 963 1012 N-linked (GlcNAc...). (Potential).
FT CARBOHYD 110 110 N-linked (GlcNAc...). (Potential).
FT CARBOHYD 180 180 N-linked (GlcNAc...). (Potential).
FT CARBOHYD 496 496 I -> V (in Ref. 2 and 5).
FT CONFLICT 124 124 M -> N (in Ref. 2).
FT CONFLICT 889 889 M -> N (in Ref. 2).
FT CONFLICT 945 945 M -> N (in Ref. 2).
SQ SEQUENCE 1043 AA; 112243 MW; ADE46133F8B77C11 CRC64;

Query Match 83.5%; Score 81; DB 1; Length 1043;

Best Local Similarity 78.9%; Pred. No. 9.5e-05;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
DB 186 EPNHLSKIAFKIIRQEPS 204

RESULT 11

DSGI_HUMAN STANDARD; PRT; 1049 AA.

AC Q02413;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Desmoglein 1 precursor (Desmosomal glycoprotein 1) (DGI) (DGI)
DE (Memphigus foliaceus antigen).
GN Name=DSG1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Keratinocytes;
RX MEDLINE=91271279; PubMed=1711210;
RA Wheeler G.N., Parker A.E., Thomas C.L., Ataliotis P., Poynter D.,
RA Arnenmann J., Rutman A.J., Pidsley S.C., Watt F.M., Rees D.A.,
RA Buxton R.S., Magee A.I.;
RT "Desmosomal glycoprotein DGI, a component of intercellular desmosome
RT junctions, is related to the cadherin family of cell adhesion
RT molecules.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:4796-4800(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Foreskin;
RX MEDLINE=92121251; PubMed=1770008;
RA Nilles L.A., Parry D.A., Powers E.E., Angst B.D., Wagner R.M.,
RA Green K.J.;
RT "Structural analysis and expression of human desmoglein: a cadherin-
RT like component of the desmosome.";
RL J. Cell Sci. 99:809-821(1991).
RN [3]
RP INVOLVEMENT IN PPKS1.
RX PubMed=10332028; DOI=10.1093/hmg/8.6.971;
RA Rickman L., Smrak D., Stevens H.P., Hunt D.M., King I.A.,
RA Bryant S.P., Eady R.A.J., Leigh I.M., Arnenmann J., Magee A.I.,
RA Kelsell D.P., Buxton R.S.;
RT "N-terminal deletion in a desmosomal cadherin causes the autosomal
RT dominant skin disease striate palmoplantar keratoderma.";
RL Hum. Mol. Genet. 8:971-976(1999).
CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved
CC in the interaction of plaque proteins and intermediate filaments
CC mediating cell-cell adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Epidermis, tongue, tonsil and esophagus.
CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
CC (Potential).
CC -!- DISEASE: Defects in DSG1 are the cause of keratosis palmoplantaris
CC striata I (PPKS1) [MIM:148700]; also known as striate palmoplantar
CC keratoderma I (SPK1). PPKS1 is an autosomal dominant disease
CC characterized by thickening of the skin on the palms and soles,
CC and longitudinal hyperkeratotic lesions on the palms, running the
CC length of each finger.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X56654; CAA39976.1; -.
DR EMBL; AF097935; AAC83817.1; -.
DR FIR; S16906; IJHUG1.
DR HSSP; P15116; INCJ.
DR Genew; HGNC:3048; DSG1.
DR MIM; 125670; -.
DR MIM; 148700; -.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin C term.
DR InterPro; IPR009123; Desmoglein.
DR InterPro; IPR009122; Desmo_cadherin.
DR Pfam; PF00028; Cadherin; 4.
DR Pfam; PF01049; Cadherin C; 1.
DR PRINTS; PR00205; CADHERIN.
DR PRINTS; PR01818; DESMOCADHERN.
DR PRINTS; PR01819; DESMOGLEIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00332; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 4.
KW Calcium-binding; Cell adhesion; Cytoskeleton; Glycoprotein; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 23 Potential.
FT PROPEP 24 49 Potential.
FT CHAIN 50 1049 Desmoglein 1.
FT DOMAIN 50 545 Extracellular (Potential).
FT TRANSMEM 546 570 Potential.
FT DOMAIN 571 1049 Cytoplasmic (Potential).
FT DOMAIN 50 158 Cadherin 1.
FT DOMAIN 159 270 Cadherin 2.
FT DOMAIN 271 385 Cadherin 3.
FT DOMAIN 386 497 Cadherin 4.
FT REPEAT 813 839 Desmoglein repeat 1.
FT REPEAT 840 869 Desmoglein repeat 2.
FT REPEAT 870 899 Desmoglein repeat 3.
FT REPEAT 900 927 Desmoglein repeat 4.
FT REPEAT 928 956 Desmoglein repeat 5.
FT DOMAIN 969 1019 Gly/Ser-rich.
FT CARBOHYD 36 36 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 110 110 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 180 180 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1049 AA; 113715 MW; EEE125655B9D6619 CRC64;
Query Match 83.5%; Score 81; DB 1; Length 1049;
Best Local Similarity 78.9%; Pred. No. 9.6e-05;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
DB 186 EPNHLSKIAFKIIRQEPS 204

RESULT 12

Q9GKQ8 PRELIMINARY; PRT; 1054 AA.

AC Q9GKQ8;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Desmoglein-1 precursor.
GN Name=DSG-1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lip;
RX MEDLINE=20304882; PubMed=10844571;
RA Muller E., Caldelari R., Levine R., Kaplan S., Baron A., RohrbachB,
RA Wyder M., Balmer V., Suter M.M.;
RT "Cloning of canine Dgl and evidence for alternative

```

RT polyadenylation."
RL J. Invest. Dermatol. 114:1211-1213(2000).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Contains 2 cadherin domains.
DR EMBL; AF005360; AAD01241.1; -.
DR HSSP; P15116; 1NCJ.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR002233; Cadherin_C_term.
DR InterPro; IPR009123; Desmoglein.
DR InterPro; IPR009122; Desmo_cadherin.
DR Pfam; PF00028; Cadherin; 4.
DR Pfam; PF01049; Cadherin_C; 1.
DR PRINTS; PR02025; CADHERIN.
DR PRINTS; PR01818; DESMOCADHERN.
DR PRINTS; PR01819; DESMOGLEIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS50268; CADHERIN_2; 2.
KW Calcium; Calcium-binding; Cell adhesion; Signal; Transmembrane.
FT SIGNAL 1 21 Potential.
FT CHAIN 50 1054 desmoglein-1.
SQ SEQUENCE 1054 AA; 113857 MW; B8690F48E5F303 CRC64;

Query Match 83.5%; Score 81; DB 2; Length 1054;
Best Local Similarity 78.9%; Pred. No. 9.6e-05;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
DB 186 EPNHLSKIAFKIIRQEPS 204
|||||:|||||:|||||:

RESULT 13
QYTQ61 PRELIMINARY; PRT; 911 AA.
AC QYTQ61
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Desmoglein lgamma.
GN Name=Dsglgamma;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=PWK;
RC MEDLINE=22518810; PubMed=12631243;
RA Kljuic A., Cristiano A.M.;
RT "A novel mouse desmosomal cadherin family member, desmoglein 1
gamma."
RL Exp. Dermatol. 12:20-29(2003).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Contains 3 cadherin domains.
DR EMBL; AV314983; AAP79926.1; -.
DR HSSP; P09803; 1FF5.
DR GO; GO:0030057; C:desmosome; IC.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR002233; Cadherin_C_term.
DR Pfam; PF00028; Cadherin; 3.
DR Pfam; PF01049; Cadherin_C; 1.
DR SMART; SM00112; CA; 3.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS00232; CADHERIN_2; 2.

```

```

DR PROSITE; PS50268; CADHERIN_2; 3.
KW Calcium-binding; Cell adhesion; Transmembrane.
SQ SEQUENCE 911 AA; 100453 MW; D456C9431BC6941C CRC64;

Query Match 77.3%; Score 75; DB 2; Length 911;
Best Local Similarity 73.7%; Pred. No. 0.00088;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
DB 186 EPNHLSKIAFKIIRQEPS 204
|||||:|||||:|||||:

RESULT 14
QYTQ60 PRELIMINARY; PRT; 911 AA.
AC QYTQ60
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Desmoglein 6.
GN Name=Dsg6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22672408;
RA Whittcock N.V.;
RT "Genetic analysis of the mouse desmoglein cluster reveals evidence for
six distinct genes: Characterization of Mouse DSG4, DSG5, and DSG6."
RL J. Invest. Dermatol. 120:970-980(2003).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Contains 3 cadherin domains.
DR EMBL; AY192159; AAP31153.1; -.
DR HSSP; P09803; 1FF5.
DR GO; GO:0030057; C:desmosome; IC.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR002233; Cadherin_C_term.
DR Pfam; PF00028; Cadherin; 3.
DR Pfam; PF01049; Cadherin_C; 1.
DR SMART; SM00112; CA; 3.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS50268; CADHERIN_2; 3.
KW Calcium-binding; Cell adhesion; Transmembrane.
SQ SEQUENCE 911 AA; 100515 MW; 2864E998106CFB6F CRC64;

Query Match 77.3%; Score 75; DB 2; Length 911;
Best Local Similarity 73.7%; Pred. No. 0.00088;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
DB 186 EPNHLSKIAFKIIRQEPS 204
|||||:|||||:|||||:

RESULT 15
QYTQ60 PRELIMINARY; PRT; 1060 AA.
AC QYTQ60
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Desmoglein lbeta.
GN Name=Dsglb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN

```

RP SEQUENCE FROM N.A.
 RC STRAIN=PWK;
 RX MEDLINE=22518809; PubMed=12631242;
 RT Pulkkinen L., Choi Y.W., Kljuic A., Uitto J., Mahoney M.G.;
 RA "Novel member of the mouse desmoglein gene family: Dsg1-beta";
 RL Exp. Dermatol. 12:11-19(2003).
 CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins
 CC (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- SIMILARITY: Contains 2 cadherin domains.
 DR EMBL; AY315940; AAP80571.1; -.
 DR HSP; P09803; 1FF5.
 DR GO; GO:0030057; C:desmosome; IDA.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000233; Cadherin_C_term.
 DR Pfam; PF00028; Cadherin; 4.
 DR Pfam; PF01049; Cadherin_C; 1.
 DR SMART; SM00112; CA; 4.
 DR PROSITE; PS00232; CADHERIN_1; 2.
 DR PROSITE; PS0268; CADHERIN_2; 2.
 KW Calcium; Calcium-binding; Cell adhesion; Transmembrane.
 SQ SEQUENCE 1060 AA; 114452 MW; F79BC83C9E474FA4 CRC64;

Query Match 77.3%; Score 75; DB 2; Length 1060;
 Best Local Similarity 73.7%; Pred. No. 0.001;
 Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EPNHLSKIAFKIVSQEPA 19
 Db 186 EPNHLSKIAFKIVSQEPA 204

Search completed: March 15, 2005, 11:53:52
 Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 11:31:32 ; Search time 168 seconds
(without alignments)
43.741 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*

Maximum Match 100*

Listing first 45 summaries

Database : A Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	614	2 AAW07908	Pemphigus
2	97	100.0	999	2 AAR30742	Aar30742 Human pem
3	97	100.0	999	5 AAU78054	AAU78054 Human des
4	97	100.0	999	6 ABR48239	ABR48239 Human bla
5	97	100.0	999	6 ABU56419	ABU56419 Lung canc
6	97	100.0	999	6 ABUS6662	ABUS6662 Lung canc
7	97	100.0	999	6 ADA83744	ADA83744 Human DSG
8	97	100.0	999	7 ADW47209	ADW47209 Desmoglei
9	97	100.0	999	7 ADN39004	ADN39004 Cancer/an
10	97	100.0	999	8 ADJ75518	ADJ75518 Marker ge
11	97	100.0	999	8 ADN04164	ADN04164 Antipsori
12	97	100.0	999	8 ABM81960	ABM81960 Tumour-as
13	97	100.0	1014	4 ABG12435	ABG12435 Novel hum
14	83	85.6	274	4 ABG12434	ABG12434 Novel hum
15	83	85.6	1191	4 ABG12436	ABG12436 Novel hum
16	81	83.5	112	3 AAY64603	AAY64603 Nonclassi
17	81	83.5	778	2 AAU15489	AAU15489 Pemphigus
18	81	83.5	1049	5 AAU78052	AAU78052 Human des
19	81	83.5	1049	8 ADQ19903	ADQ19903 Human sof
20	81	83.5	1049	8 ADQ14316	ADQ14316 Human des
21	81	83.5	1049	8 ADR41698	ADR41698 Human des
22	81	83.5	1049	8 ADP56044	ADP56044 Human pro
23	81	83.5	1054	8 ADR66714	ADR66714 Human pro
24	81	83.5	1054	8 ADR66372	ADR66372 Human pro
25	71	73.2	15	2 AAW04843	AAW04843 Self epit

26	71	73.2	15	2 AAW64815	AAW64815 Desmoglei
27	71	73.2	15	2 AAW78814	AAW78814 Desmoglei
28	71	73.2	15	3 AAB33625	AAB33625 MHC class
29	71	73.2	15	4 AAG93721	AAG93721 Human des
30	71	73.2	15	5 AAO17033	AAO17033 Desmoglei
31	71	73.2	15	6 ABU96577	ABU96577 MHC class
32	71	73.2	15	8 ADQ14318	ADQ14318 Human des
33	71	73.2	15	8 ADR41700	ADR41700 Desmoglei
34	71	73.2	15	8 ADS14310	ADS14310 Desmoglei
35	71	73.2	113	3 AAY64604	AAY64604 Nonclassi
36	71	73.2	263	2 AAW13010	AAW13010 Segment o
37	71	73.2	560	2 AAW13009	AAW13009 Segment o
38	71	73.2	940	4 AAM39436	AAM39436 Human pol
39	71	73.2	1117	5 AAU78053	AAU78053 Human des
40	71	73.2	1117	8 ADR99232	ADR99232 Clone IMA
41	71	73.2	1118	5 AAE20266	AAE20266 Human lun
42	71	73.2	1120	4 ABG10284	ABG10284 Novel hum
43	71	73.2	1121	4 AAM41222	AAM41222 Human pol
44	52	53.6	747	7 ADD14035	ADD14035 Human src
45	52	53.6	761	8 ADP29437	ADP29437 Human sec

ALIGNMENTS

RESULT 1

AAW07908

ID AAW07908 standard; protein; 614 AA.

XX

AC AAW07908;

XX

DT 29-JAN-1997 (first entry)

XX

DE Pemphigus vulgaris antigen protein extracellular region.

XX

KW Autoantibody; immunoglobulin G; IgG1; fusion protein; diagnosis;

KW treatment; pemphigus vulgaris; PV; bulla; blister; skin disease;

KW dermatology.

XX

OS Homo sapiens.

XX

PN JP08188540-A.

XX

PD 23-JUL-1996.

XX

PF 30-JUN-1995; 95JP-00165632.

XX

PR 30-JUN-1994; 94JP-00173291.

XX

PA (NISH/) NISHUKAWA T.

XX

DR WPI; 1996-388562/39.

XX

PT Fused protein recognised by Pemphigus vulgaris auto-antibody - useful to

PT treat and diagnose P. vulgaris related diseases.

XX

Claim 1; Page 7-9; 9pp; Japanese.

XX

AAW07908 represents the human pemphigus vulgaris (PV) antigen extracellular region. The PV antigen is produced in patients with pemphigus vulgaris resulting in autoimmune disease. PV is a rare relapsing disease causing suprabasal, intra-epidermal bullae (vesicles) of the skin and mucous membranes, which is fatal if untreated. The PV antigen was fused to a human IgG1 hinge region and the resulting fusion protein is useful to treat or diagnose pemphigus vulgaris

XX

SQ Sequence 614 AA;

Query Match

Best Local Similarity 100.0%; Score 97; DB 2; Length 614;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19

```

Db      185 EPNHLSKIAFKIVSQEPA 203
|||||
RESULT 2
AAR30742
ID      AAR30742 standard; protein; 999 AA.
XX
AC      AAR30742;
XX
DT      17-DEC-2001 (revised)
DT      14-JUN-1993 (first entry)
XX
DE      Human pemphigus vulgaris 130kD antigen.
XX
KW      Pemphigus vulgaris; skin disease; autoantibodies;
KW      keratinocyte cell surface antigen; glycoprotein; cell adhesion.
XX
OS      Homo sapiens.
XX
PN      USN7798918-N.
XX
PD      15-DEC-1992.
XX
PF      27-NOV-1991; 91US-00798918.
XX
PR      27-NOV-1991; 91US-00798918.
XX
PA      (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI      Stanley JR, Amagai M, Klaus-Kovtun V;
XX
DR      WPI; 1993-067436/08.
DR      N-PSDB; AAQ35992.
XX
XX      DNA encoding pemphigus vulgaris antigen - useful in proteins for
PT      diagnostic and therapeutic uses.
XX
PS      Disclosure; Fig 7; 50pp; English.
XX
CC      This sequence is the pemphigus vulgaris 130kD antigen. The protein and
CC      its encoding DNA may be used in the diagnosis and treatment of pemphigus
CC      vulgaris. It is thought that the antigen may be a cell adhesion molecule.
CC      (Note: Revised entry submitted to correct the patent number format of US
CC      Government-owned NTIS applications to prevent clashes with ongoing US
CC      granted patent numbers. For further information please visit the Derwent
CC      web site at www.derwent.com/dwpi/updates/ntis\_us.html.)
XX
SQ      Sequence 999 AA;

Query Match      100.0%; Score 97; DB 2; Length 999;
Best Local Similarity 100.0%; Pred. No. 7.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EPNHLSKIAFKIVSQEPA 19
|||||
Db      186 EPNHLSKIAFKIVSQEPA 204

RESULT 3
AAU78054
ID      AAU78054 standard; protein; 999 AA.
XX
AC      AAU78054;
XX
DT      05-JUN-2002 (first entry)
XX
DE      Human desmoglein 3 (pemphigus vulgaris antigen) protein sequence.
XX
KW      Human; desmoglein; stratified squamous epithelia; noxious substance;
KW      extracellular domain; junctional protein; luminal damage; heartburn;
KW      laryngitis; gastric influx; gastroesophageal reflux disease; GERD;
KW      sore throat; pemphigus vulgaris antigen.
XX

```

```

XX      Homo sapiens.
XX
FH      Key
FT      Domain
FT      Location/Qualifiers
FT      50..615
FT      /label= Extracellular_domain
FT      52..157
FT      /note= "Specifically claimed in claim 8"
XX
PN      WO200210767-A2.
XX
PD      07-FEB-2002.
XX
PF      26-JUL-2001; 2001WO-US023717.
XX
PR      28-JUL-2000; 2000US-00626196.
XX
PA      (TULA ) TULANE EDUCATIONAL FUND.
XX
PI      Tobey NA, Orlando RC;
XX
DR      WPI; 2002-257387/30.
XX
PT      Identifying compounds that protect stratified squamous epithelial tissue
PT      against injury by noxious substances, by determining interaction between
PT      a test compound and an E-cadherin extracellular domain.
XX
PS      Claim 7; Page 53-55; 62pp; English.
XX
CC      The present invention relates to a new method of assaying for a compound
CC      which may protect stratified squamous epithelia from damage by a noxious
CC      substance. The method of the invention involves determining the level or
CC      presence of an interaction between the test compound and a polypeptide
CC      sequence comprising a portion of the extracellular domain of the
CC      junctional protein E-cadherin or a related polypeptide sequence. The
CC      method is useful for identifying compounds, which may protect stratified
CC      squamous epithelial tissue against injury by noxious substances. In
CC      particular, the method is useful for identifying compounds for protective
CC      effects against luminal damaging compounds or compounds that may be
CC      useful in treating conditions associated with gastric influx e.g.
CC      gastroesophageal reflux disease (GERD), or in treating sore throat,
CC      heartburn or laryngitis. The present amino acid sequence represents the
CC      human desmoglein 3 protein of the invention. This protein is also known
CC      as pemphigus vulgaris antigen
XX
SQ      Sequence 999 AA;

Query Match      100.0%; Score 97; DB 5; Length 999;
Best Local Similarity 100.0%; Pred. No. 7.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EPNHLSKIAFKIVSQEPA 19
|||||
Db      186 EPNHLSKIAFKIVSQEPA 204

RESULT 4
ABR48239
ID      ABR48239 standard; protein; 999 AA.
XX
AC      ABR48239;
XX
DT      12-JUN-2003 (first entry)
XX
DE      Human bladder cancer associated protein sequence SEQ ID NO:199.
XX
KW      Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS      Homo sapiens.
XX
PN      WO2003003906-A2.
XX
PD      16-JAN-2003.

```

```

XX 03-JUL-2002; 2002WO-US021338.
XX
XX 03-JUL-2001; 2001US-0302814P.
XX
XX 03-AUG-2001; 2001US-0310099P.
XX
XX 08-NOV-2001; 2001US-0343705P.
XX
XX 13-NOV-2001; 2001US-0350666P.
XX
XX 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Mack DH, Aziz N;
XX
XX WPI: 2003-093161/08.
XX
XX N-PSDB; ACC51055.
XX
XX Detecting a bladder cancer-associated transcript in a cell from a
XX patient, comprises contacting a biological sample from the patient with a
XX bladder cancer-associated polynucleotide or antibody.
XX
XX Claim 10; Page 299; 307pp; English.
XX
XX The present invention describes a method for detecting a bladder cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with a polynucleotide
XX that selectively hybridizes to a sequence that is 80 % identical to a
XX table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
XX encode the human bladder cancer-associated proteins given in ABR48146 to
XX ABR48242). Bladder cancer-associated sequences from the present invention
XX have cytostatic activities, and can be used in antisense gene therapy and
XX in vaccine production. The method can be used for detecting a bladder
XX cancer-associated transcript in a cell from a patient. The method is
XX useful in diagnosing or treating bladder cancer and in screening for
XX compounds that modulate bladder cancer, such as hormones or antibodies.
XX The nucleic acid molecules from the present invention may be used in
XX various screening and diagnostic methods, and for gene therapy, vaccine
XX and/or antisense/inhibition applications
XX
XX Sequence 999 AA;
XX
XX Query Match 100.0%; Score 97; DB 6; Length 999;
XX Best Local Similarity 100.0%; Pred. No. 7.4e-08; Mismatches 0; Indels 0; Gaps 0;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EPNHLSKIAPKIVSQEPA 19
XX |||||
XX Db 186 EPNHLSKIAPKIVSQEPA 204
XX
XX RESULT 5
XX ABUS6419
XX ID ABUS6419 standard; protein; 999 AA.
XX
XX AC ABUS6419;
XX
XX DT 02-APR-2003 (first entry)
XX
XX DE Lung cancer-associated polypeptide #12.
XX
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
XX OS Unidentified.
XX
XX PN WO200286443-A2.
XX
XX PD 31-OCT-2002.
XX
XX PF 18-APR-2002; 2002WO-US012476.
XX
XX
XX 03-JUL-2002; 2002WO-US021338.
XX
XX 03-JUL-2001; 2001US-0302814P.
XX
XX 03-AUG-2001; 2001US-0310099P.
XX
XX 08-NOV-2001; 2001US-0343705P.
XX
XX 13-NOV-2001; 2001US-0350666P.
XX
XX 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Aziz N, Murray R;
XX
XX WPI: 2003-093161/08.
XX
XX N-PSDB; ABX76135.
XX
XX Detecting a lung cancer-associated transcript in a cell from a patient
XX for treating lung cancer, by contacting a biological sample from the
XX patient with a polynucleotide that exhibits increased or decreased
XX expression in lung cancer.
XX
XX Claim 27; Page 197; 453pp; English.
XX
XX The invention relates to a method for detecting a lung cancer-associated
XX transcript in a cell from a patient, comprising contacting a biological
XX sample from the patient with a polynucleotide that selectively hybridizes
XX to a sequence that is at least 80 % identical to a gene that exhibits
XX increased or decreased expression in lung cancer samples. Lung cancer-
XX associated polynucleotides and polypeptides are used for identifying a
XX compound that modulates a lung cancer-associated polypeptide, for
XX inhibiting proliferation of a lung cancer-associated cell to treat lung
XX cancer in a patient and for treating a mammal having lung cancer by
XX administering a modulatory compound identified. The methods are useful
XX for treating lung cancer, such as small cell lung cancer, non-small cell
XX lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX for diagnostic purposes and as targets for screening for therapeutic
XX compounds that modulate lung cancer, such as antibodies. Sequences
XX ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
XX invention
XX
XX Sequence 999 AA;
XX
XX Query Match 100.0%; Score 97; DB 6; Length 999;
XX Best Local Similarity 100.0%; Pred. No. 7.4e-08;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EPNHLSKIAPKIVSQEPA 19
XX |||||
XX Db 186 EPNHLSKIAPKIVSQEPA 204
XX
XX RESULT 6
XX ABUS6662
XX ID ABUS6662 standard; protein; 999 AA.
XX
XX AC ABUS6662;
XX
XX DT 02-APR-2003 (first entry)
XX
XX DE Lung cancer-associated polypeptide #255.
XX
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
XX OS Unidentified.
XX
XX PN WO200286443-A2.
XX
XX PD 31-OCT-2002.
XX
XX PF 18-APR-2002; 2002WO-US012476.
XX
XX

```

```

XX PF 18-APR-2002; 2002WO-US012476.
XX PR 18-APR-2001; 2001US-0284770P.
XX PR 10-MAY-2001; 2001US-0290492P.
XX PR 09-NOV-2001; 2001US-0339245P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 29-NOV-2001; 2001US-0334370P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX DR WPI; 2003-093161/08.
XX DR N-PSDB; ABX76391.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
XX PT for treating lung cancer, by contacting a biological sample from the
XX PT patient with a polynucleotide that exhibits increased or decreased
XX PT expression in lung cancer.
XX PS Claim 27; Page 385; 453pp; English.
XX CC The invention relates to a method for detecting a lung cancer-associated
XX CC transcript in a cell from a patient, comprising contacting a biological
XX CC sample from the patient with a polynucleotide that selectively hybridizes
XX CC to a sequence that is at least 80 % identical to a gene that exhibits
XX CC increased or decreased expression in lung cancer samples. Lung cancer-
XX CC associated polynucleotides and polypeptides are used for identifying a
XX CC compound that modulates a lung cancer-associated polypeptide, for
XX CC inhibiting proliferation of a lung cancer-associated cell to treat lung
XX CC cancer in a patient and for treating a mammal having lung cancer by
XX CC administering a modulatory compound identified. The methods are useful
XX CC for treating lung cancer, such as small cell lung cancer, non-small cell
XX CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX CC for diagnostic purposes and as targets for screening for therapeutic
XX CC compounds that modulate lung cancer, such as antibodies. Sequences
XX CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
XX CC invention
XX CC Sequence 999 AA;
XX SQ
Query Match 100.0%; Score 97; DB 6; Length 999;
Best Local Similarity 100.0%; Pred. No. 7.4e-08; Mismatches 0; Gaps 0;
Matches 19; Conservative 0; Indels 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
Db 186 EPNHLSKIAFKIVSQEPA 204

RESULT 7
ADA83744
ID ADA83744 standard; protein; 999 AA.
XX AC ADA83744;
XX DT 20-NOV-2003 (first entry)
XX DE Human DSG3 protein.
XX KW human; marker; expressed sequence tag; EST; arabidopsis; tumour;
XX KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
XX KW vaccine.
XX OS Homo sapiens.
XX PN WO2002103028-A2.
XX XX

```

```

PD 27-DEC-2002.
XX PF 30-MAY-2002; 2002WO-IB004189.
XX PR 30-MAY-2001; 2001US-0293399P.
XX PR 22-OCT-2001; 2001US-0330457P.
XX PR 19-FEB-2002; 2002US-0357144P.
XX PA (BIOM-) BIOMEDICAL CENT.
XX PI Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LI;
XX DR WPI; 2003-175241/17.
XX DR N-PSDB; ADA83743.
XX PT Determining if a nucleic acid is a marker for a phenotype/cell type of
XX PT interest, by global comparison of expressed sequence tags known to be
XX PT expressed in the phenotype/cell type with all ESTs expressed in normal
XX PT tissue.
XX PS Claim 29; Page 96-101; 516pp; English.
XX CC The invention relates to a novel method for determining if a nucleic acid
XX CC is a marker for a predetermined phenotype/cell type of interest from a
XX CC biological species. The method comprises performing a global comparison
XX CC of a group of expressed sequence tags (ESTs) known to be expressed in the
XX CC phenotype/cell type of interest with all ESTs expressed in normal tissue
XX CC in order to identify ESTs that are preferentially expressed in the
XX CC phenotype/cell type of interest. A method of the invention is useful for
XX CC determining whether a nucleic acid is a marker for a predetermined
XX CC phenotype or cell type of interest from a biological species, preferably
XX CC Arabidopsis or human. The cell type of interest is an abnormal cell such
XX CC as a tumour cell, and the predetermined phenotype is a stress-induced
XX CC phenotype such as hyperosmotic stress or high salt conditions. A method
XX CC of the invention is also useful for determining the progression of colon
XX CC cancer in a human, for detecting a tumour cell, and for regulating or
XX CC preventing the growth of a tumour cell. An antibody of the invention is
XX CC useful for detecting the absence or presence of peptides encoded by an
XX CC tumour-associated markers. A polypeptide of the invention is useful as an
XX CC immunogen for vaccinating an animal. The present sequence represents a
XX CC tumour-associated antigen of the invention.
XX CC Sequence 999 AA;
XX SQ
Query Match 100.0%; Score 97; DB 6; Length 999;
Best Local Similarity 100.0%; Pred. No. 7.4e-08; Mismatches 0; Gaps 0;
Matches 19; Conservative 0; Indels 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
Db 186 EPNHLSKIAFKIVSQEPA 204

RESULT 8
ADA47209
ID ADA47209 standard; protein; 999 AA.
XX AC ADA47209;
XX DT 03-JUN-2004 (first entry)
XX DE Desmoglein 3 variant like NOVX 10a protein.
XX KW NOVX; cytostatic; gene therapy; vaccine; cancer; chromosome mapping.
XX OS Unidentified.
XX PN WO2003083039-A2.
XX XX
XX PD 09-OCT-2003.
XX PF 03-JUL-2002; 2002WO-US021485.
XX XX

```

05-JUL-2001; 2001US-0303046P.
09-JUL-2001; 2001US-0303282P.
11-JUL-2001; 2001US-0304502P.
12-JUL-2001; 2001US-0305011P.
13-JUL-2001; 2001US-0305262P.
16-JUL-2001; 2001US-0305673P.
17-JUL-2001; 2001US-0306085P.
24-JUL-2001; 2001US-0307536P.
27-JUL-2001; 2001US-0308228P.
30-JUL-2001; 2001US-0308877P.
14-AUG-2001; 2001US-0312203P.
17-SEP-2001; 2001US-0322640P.
19-SEP-2001; 2001US-0323484P.
21-SEP-2001; 2001US-0323821P.
21-SEP-2001; 2001US-0323948P.
25-SEP-2001; 2001US-0324711P.
09-OCT-2001; 2001US-0327893P.
21-NOV-2001; 2001US-0331768P.
21-FEB-2002; 2002US-0359191P.
22-FEB-2002; 2002US-0358939P.
28-FEB-2002; 2002US-0360923P.
01-MAR-2002; 2002US-0360830P.
01-MAR-2002; 2002US-0361178P.
05-MAR-2002; 2002US-0361748P.
12-MAR-2002; 2002US-0363429P.
12-MAR-2002; 2002US-0363683P.
12-APR-2002; 2002US-0372141P.
16-APR-2002; 2002US-0372967P.
16-APR-2002; 2002US-0373051P.
16-APR-2002; 2002US-0373063P.
17-APR-2002; 2002US-0373280P.
19-APR-2002; 2002US-0373287P.
19-APR-2002; 2002US-0373981P.
02-JUL-2002; 2002US-00187975.
(CURA-) CURAGEN CORP.
Li L, Shenoy SG, Patturajan M, Ellerman K, Gorman L, Zhong M;
Catterton E, Spytek KA, Miller CB, Edinger SR, Hjal T, Gerlach VL;
Shinketsa R, Taupier RJ, Anderson DW, Guo X, Baumgartner JC;
Padigaru M, Feyman JA, Smithson G, Casman SJ, Voss EZ, Boldog FL;
Pena CEA, Chapoval A, Rastelli L, Kekuda R, Vernet CM;
WPI: 2003-812538/76.
N-PSDB; ADM47208.
New NOVX polypeptide, useful for preparing a composition for treating or preventing e.g. cancer or for chromosome mapping.
Claim 2; SEQ ID NO 42; 433pp; English.
The invention relates to a novel isolated polypeptide, designated NOVX. The novel polypeptide comprises a sequence comprising 109-1671 amino acids, or its mature form; a sequence that is at least 95% identical to the 109-1671 amino acid polypeptide; or a sequence comprising one or more conservative substitutions in the 109-1671 amino acid polypeptide. The invention further comprises: a composition; a kit comprising the composition; a method for determining the presence or amount of the polypeptide or nucleic acid molecule in a sample; determining the presence of, or predisposition to, a disease associated with the altered levels of nucleic acid or of expression of the polypeptide in a first mammalian subject; identification of an agent that binds to the polypeptide; identification of a potential therapeutic agent for treating a pathology related to aberrant expression or physiological interactions of the polypeptide; a method of screening for a modulator of activity or latency of, or predisposition to, a pathology associated with the polypeptide; a method for modulating the activity of the polypeptide; treating or preventing a pathology associated with the polypeptide; treating a pathological state in a mammal; an isolated nucleic acid molecule; a vector comprising the nucleic acid molecule; a cell comprising the vector; an antibody that immunospecifically binds to the polypeptide; and a method for producing the polypeptide. The NOVX polypeptide and its encoding nucleic acid have cytostatic activity. The

CC NOVX polynucleotide can be used in gene therapy to treat disorders. The CC NOVX polypeptide can be used to create a vaccine. The polypeptide is CC useful for preparing a composition for treating or preventing a CC pathological state in a mammal, e.g., cancer, or for chromosome mapping. CC This sequence represents a NOVX polypeptide of the invention.
XX
SQ Sequence 999 AA;
Query Match 100.0%; Score 97; DB 7; Length 999;
Best Local Similarity 100.0%; Pred. No. 7.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPHLNSKIAFKIVSQEPA 19
DB 186 EPHLNSKIAFKIVSQEPA 204
RESULT 9
ADN39004
ID ADN39004 standard; protein; 999 AA.
XX
AC ADN39004;
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:322.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-039775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI: 2003-468649/44.
DR N-PSDB; ADN39003.
XX
PT Determining the presence or absence of a pathological cell in a patient, PT useful for diagnosing, prognosing or treating cancer, comprises detecting a nucleic acid in a biological sample.
PT

XX Claim 12; SEQ ID NO 322; 1385pp; English.

PS The invention relates to nucleic acids and proteins (ADN38693-ADN40064)

CC whose expression is upregulated or downregulated in specific cancers or

CC other diseases such as angiogenic or fibrotic disorders, and to methods

CC of determining the presence or absence of a pathological cell in a

CC patient by detecting a nucleic acid at least 80% identical to those of

CC the invention or by detecting a polypeptide of the invention. The

CC invention also relates to expression vectors and host cells comprising a

CC nucleic acid of the invention; antibodies which specifically bind a

CC polypeptide of the invention; use of such antibodies for drug targeting;

CC and methods of screening for modulators of activity or expression of the

CC polypeptides and nucleic acids. The nucleic acids, polypeptides,

CC antibodies and methods are useful for diagnosing, prognosing and treating

CC cancer and other conditions such as psoriasis, ischaemia, heart disease,

CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal

CC neovascularisation syndromes, scarring and uterine fibroids. They may

CC also be useful in wound healing and in contraception. The present

CC sequence represents a polypeptide of the invention.

XX

SQ Sequence 999 AA;

Query Match 100.0%; Score 97; DB 7; Length 999;

Best Local Similarity 100.0%; Pred. No. 7.4e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19

Db 186 EPNHLSKIAFKIVSQEPA 204

|||||

RESULT 10

ADJ75518

ID ADJ75518 standard; protein; 999 AA.

XX

AC ADJ75518;

DT 20-MAY-2004 (first entry)

XX

DE Marker gene related amino acid sequence SEQ ID NO:770.

XX

XX bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

KW gene therapy; marker.

XX

OS Homo sapiens.

XX

PN EF1394274-A2.

XX

PD 03-MAR-2004.

XX

PF 04-AUG-2003; 2003EP-00254857.

XX

PR 06-AUG-2002; 2002JP-00229312.

PR 20-MAR-2003; 2003JP-00077212.

XX

PA (GENO-) GENOX RES INC.

XX

PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX

XX WPI; 2004-193155/19.

DR

XX

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by

PT comparing the expression level of a marker gene in a biological sample

PT from a subject with the expression level of the gene in a sample from a

PT healthy subject.

XX

XX Example 11; SEQ ID NO 770; 241pp; English.

PS

XX The present invention describes a method of testing for bronchial asthma

CC or chronic obstructive pulmonary disease. The method comprises

CC determining the expression level of a marker gene in a biological sample

CC from a subject, comparing the expression level determined with the

CC expression level of the marker gene in a biological sample from a healthy

CC subject, and judging whether the subject has bronchial asthma or chronic

CC obstructive pulmonary disease. The marker gene comprises: (a) a group of

CC genes (S1) whose expression levels increase when respiratory epithelial

CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)

CC whose expression levels decrease when respiratory epithelial cells are

CC stimulated with interleukin-13. Also described: (1) a reagent (I) for

CC testing for bronchial asthma or chronic obstructive pulmonary disease;

CC (2) a kit for screening for a candidate compound for a therapeutic agent

CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)

CC an animal model for bronchial asthma or chronic obstructive pulmonary

CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a

CC method for producing an animal model for bronchial asthma or chronic

CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial

CC asthma or chronic obstructive pulmonary disease, comprising a portion of

CC a marker gene or an antisense nucleic acid corresponding to a portion of

CC the marker gene, a ribozyme, a polynucleotide that suppresses the

CC expression of the gene through an RNAi effect or an antibody recognising

CC a protein encoded by a marker gene; and (7) a DNA chip for testing for

CC bronchial asthma or a chronic obstructive pulmonary disease, on which a

CC probe has been immobilised to assay a marker gene. (I) has respiratory

CC and antiasthmatic activities, and can be used in gene therapy. The method

CC is useful for testing for or screening for a therapeutic agent for

CC bronchial asthma or chronic obstructive pulmonary disease. The present

CC sequence is used in the exemplification of the present invention.

XX

SQ Sequence 999 AA;

Query Match 100.0%; Score 97; DB 8; Length 999;

Best Local Similarity 100.0%; Pred. No. 7.4e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19

Db 186 EPNHLSKIAFKIVSQEPA 204

|||||

RESULT 11

ADN04164

ID ADN04164 standard; protein; 999 AA.

XX

AC ADN04164;

XX

DT 01-JUL-2004 (first entry)

XX

DE Antipsoriatic protein sequence #277.

XX

XX antipsoriatic; gene therapy; psoriasis; diagnosis.

XX

OS Homo sapiens.

XX

PN WO2004028479-A2.

XX

PD 08-APR-2004.

XX

PF 25-SEP-2003; 2003WO-US030907.

XX

PR 25-SEP-2002; 2002US-0414006P.

XX

PA (GETH) GENENTECH INC.

XX

PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

XX

XX Wu TD;

PI

XX WPI; 2004-305105/28.

DR

XX N-PSDB; ADN04163.

DR

XX New PRO nucleic acid or polypeptide, useful for preparing a

PT pharmaceutical composition for diagnosing or treating psoriasis in a

PT mammal.

XX

XX Claim 9; SEQ ID NO 558; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX
SQ Sequence 999 AA;

Query Match 100.0%; Score 97; DB 8; Length 999;
Best Local Similarity 100.0%; Pred. No. 7.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps

Qy 1 EPNHLSKIAFKIVSOEPA 19
| | | | | | | | | | | | | | | | | | | | |
Db 186 EPNHLSKIAFKIVSOEPA 204

RESULT 12
ABM81960
ID ABM81960 standard; protein; 999 AA.
XX AC - ABM81960;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO2267, SEQ:5051.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
XX WO2004030615-A2.
XX
XX 15-APR-2004.
PD
XX 29-SEP-2003; 2003WO-US028547.
XX
XX 02-OCT-2002; 2002US-0414971P.
XX
XX (GETH) GENENTECH INC.
XX
XX Wu TD, Zhang Z, Zhou Y;
XX
XX WPI; 2004-347921/32.
XX
XX N-PSDB; ACN40334.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
XX Claim 12; SEQ ID NO 5051; 7273pp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with

CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention

XX
SQ Sequence 999 AA;

Query Match 100.0%; Score 97; DB 8; Length 999;
Best Local Similarity 100.0%; Pred. No. 7.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps

Qy 1 EPNHLSKIAFKIVSQEPA 19
|||||
Db 186 EPNHLSKIAFKIVSQEPA 204
|||||

RESULT 13
ABG12435 ID ABG12435 standard; protein; 1014 AA.
XX AC ABG12435;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #12426.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN W0200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX FR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX N-PSDB; AAS76622.
XX WPI; 2001-639362/73.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX
PS Claim 20; SEQ ID NO 42794; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1014 AA;
Query Match 100.0%; Score 97; DB 4; Length 1014;
Best Local Similarity 100.0%; Pred. No. 7.6e-08; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPNHLNSKIATKIVSQEPA 19
|||||
Db 201 EPNHLNSKIATKIVSQEPA 219
RESULT 14
ABG12434
ID ABG12434 standard; protein; 274 AA.
XX
AC ABG12434;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #12425.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS76621.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 42793; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 274 AA;
Query Match 85.6%; Score 83; DB 4; Length 274;
Best Local Similarity 84.2%; Pred. No. 4.7e-06;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 EPNHLNSKIATKIVSQEPA 19
|||||
Db 166 EPNHLNSKIATKIVSQEPS 184
RESULT 15
ABG12436
ID ABG12436 standard; protein; 1191 AA.
XX
AC ABG12436;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #12427.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS76623.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 42795; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1191 AA;

Query Match 85.6%; Score 83; DB 4; Length 1191;
Best Local Similarity 84.2%; Pred. NO. 2.8e-05;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19

DB 1074 EENHLSKIAYKIVSQEPS 1092

Search completed: March 15, 2005, 11:50:52
Job time : 171 secs

This Page Blank (uspto)

```

Query Match          100.0%; Score 97; DB 14; Length 999;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EPNHLSKIAFKIVSQEPA 19
        |||||||
Db      186 EPNHLSKIAFKIVSQEPA 204

RESULT 2
US-10-187-975-42
; Sequence 42, Application US/10187975
; Publication No. US20030224982A1
; GENERAL INFORMATION:

```

us-10-799-005a-1.open.rapb

Tue Mar 15 12:10:22 2005

```
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh
; APPLICANT: Patturajan, Meera
; APPLICANT: Ellerman, Karen
; APPLICANT: Gorman, Linda
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Spytek, Kimberly
; APPLICANT: Miller, Charles
; APPLICANT: Edinger, Shlomit
; APPLICANT: Hjal, Tord
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shinkets, Richard
; APPLICANT: Taupier, Raymond J. Jr.
; APPLICANT: Anderson, David
; APPLICANT: Guo, Xiaojia
; APPLICANT: Baumgartner, Jason
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennnda
; APPLICANT: Casman, Stacie
; APPLICANT: Voss, Edward
; APPLICANT: Boldog, Ferenc
; APPLICANT: Pena, Carol
; APPLICANT: Chapoval, Andrei
; APPLICANT: Rastelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Verntee, Corine
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: 21402-397A
; CURRENT APPLICATION NUMBER: US/10/187,975
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303,046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/303,828
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304,502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305,011
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/305,262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/305,673
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/306,085
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 60/307,536
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/308,228
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/308,877
; PRIOR FILING DATE: 2001-07-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 288
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 42
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-187-975-42

Query Match 100.0%; Score 97; DB 15; Length 999;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
Db 186 EPNHLSKIAPKIVSQEPA 204

RESULT 3
US-10-295-027-322
```

```
; Sequence 322, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 322
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-322

Query Match 100.0%; Score 97; DB 15; Length 999;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
Db 186 EPNHLSKIAPKIVSQEPA 204

RESULT 4
US-10-188-832-199
; Sequence 199, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
```

Query Match	Best Local Similarity	Score	DB	Length	Mismatches	Indels	Gaps
Query Match	Best Local Similarity	Score	DB	Length	Mismatches	Indels	Gaps
83.5%	78.9%	81	15	112	3	1	0
Matches	Conservative						
15	3						
QY	1 EPNHLSKIAPKIVSQEPA 19						
DB	29 EPNHLSKIAPKIVSQEPA 47						
RESULT 7							
US-09-909-460-24							
Sequence 24, Application US/09909460							
Publication No. US20020182258A1							
GENERAL INFORMATION:							
APPLICANT: Lunsford, Lynn B.							
APPLICANT: Putnam, David							
APPLICANT: Hedley, Mary Lynn							
TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF NUCLEIC							
TITLE OF INVENTION: ACID							
FILE REFERENCE: 08191/014001							
CURRENT APPLICATION NUMBER: US/09/909,460							
CURRENT FILING DATE: 2001-07-18							
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/321,346							
PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27							
NUMBER OF SEQ ID NOS: 114							
SOFTWARE: FastSeq for Windows Version 3.0							
SEQ ID NO 24							
LENGTH: 15							
TYPE: PRT							
ORGANISM: Homo sapiens							
US-09-909-460-24							
Query Match	Best Local Similarity	Score	DB	Length	Mismatches	Indels	Gaps
73.2%	100.0%	71	9	15	0	0	0
Matches	Conservative						
15	0						
QY	5 LNSKIAPKIVSQEPA 19						
DB	1 LNSKIAPKIVSQEPA 15						
RESULT 8							
US-09-872-836-24							
Sequence 24, Application US/09872836							
Publication No. US20040142475A1							
GENERAL INFORMATION:							
APPLICANT: Barman, Shikha P.							
APPLICANT: McKeever, Una							
APPLICANT: Hedley, Mary Lynn							
TITLE OF INVENTION: DELIVERY SYSTEMS FOR BIOACTIVE AGENTS							
FILE REFERENCE: 08191-018001							
CURRENT APPLICATION NUMBER: US/09/872,836							
CURRENT FILING DATE: 2001-06-01							
PRIOR APPLICATION NUMBER: US 60/208,830							
PRIOR FILING DATE: 2000-06-02							
NUMBER OF SEQ ID NOS: 120							
SOFTWARE: FastSeq for Windows Version 4.0							
SEQ ID NO 24							
LENGTH: 15							
TYPE: PRT							
ORGANISM: Homo sapiens							
US-09-872-836-24							
Query Match	Best Local Similarity	Score	DB	Length	Mismatches	Indels	Gaps
73.2%	100.0%	71	11	15	0	0	0
Matches	Conservative						
15	0						
QY	5 LNSKIAPKIVSQEPA 19						
DB	1 LNSKIAPKIVSQEPA 15						

; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-395-032-32

Query Match 73.2%; Score 71; DB 15; Length 113;
Best Local Similarity 73.7%; Pred. No. 0.00053;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
||| ||||| :||| |||
Db 29 EPNTLSKISYRIVSLEPA 47

RESULT 12
US-10-788-792-238
; Sequence 238, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10/788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 238
; LENGTH: 1117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-788-792-238

Query Match 73.2%; Score 71; DB 16; Length 1117;
Best Local Similarity 73.7%; Pred. No. 0.007;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
||| ||||| :||| |||
Db 187 EPNTLSKISYRIVSLEPA 205

RESULT 13
US-09-909-567B-42
; Sequence 42, Application US/09909567B
; Publication No. US20030022257A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto A.
; APPLICANT: Nair, Manoj
; APPLICANT: Chen, Seiyu
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes
; FILE REFERENCE: DEX-0214
; CURRENT APPLICATION NUMBER: US/09/909,567B
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,834
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 1118
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-909-567B-42-

US-10-758-970-24
; Sequence 24, Application US/10758970
; Publication No. US20050037086A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; APPLICANT: Hsu, Yung-fueh
; APPLICANT: Tyo, Michael
; TITLE OF INVENTION: CONTINUOUS-FLOW METHOD FOR PREPARING MICROPARTICLES
; FILE REFERENCE: 08191-012001
; CURRENT APPLICATION NUMBER: US/10/758,970
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/715,708A
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 60/166,516
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-758-970-24

Query Match 73.2%; Score 71; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
||| ||||| ||||| |||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 10
US-10-006-869-32
; Sequence 32, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-869-32

Query Match 73.2%; Score 71; DB 14; Length 113;
Best Local Similarity 73.7%; Pred. No. 0.00053;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
||| ||||| :||| |||
Db 29 EPNTLSKISYRIVSLEPA 47

RESULT 11
US-10-395-032-32
; Sequence 32, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL

```
Query Match      73.2%; Score 71; DB 10; Length 1118;
Best Local Similarity 73.7%; Pred. No. 0.007;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EPNHLNSKIAFKIVSQEPA 19
Db 188 EPTLNSKISYRIVSLEPA 206

RESULT 14
US-09-746-491-43
; Sequence 43, Application US/09746491
; Patent No. US20020137202A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20020137202A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-621
; CURRENT APPLICATION NUMBER: US/09/746,491
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: USSN 60/171,329
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-746-491-43

Query Match      53.6%; Score 52; DB 9; Length 413;
Best Local Similarity 47.4%; Pred. No. 3.5;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 1 EPNHLNSKIAFKIVSQEPA 19
Db 53 DPNALNGMLRYRIVSQAPS 71

RESULT 15
US-10-482-029-283
; Sequence 283, Application US/10482029
; Publication No. US20050037445A1
; GENERAL INFORMATION:
; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482,029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 283
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-482-029-283

Query Match      53.6%; Score 52; DB 17; Length 645;
Best Local Similarity 47.4%; Pred. No. 5.7;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 1 EPNHLNSKIAFKIVSQEPA 19
Db 296 DPNALNGMLRYRIVSQAPS 314

Search completed: March 15, 2005, 12:06:52
Job time : 140 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 11:44:53 ; Search time 46 Seconds
(without alignments)
30.833 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	1014	4	US-09-949-016-8229
2	81	83.5	112	3	US-09-187-859-31
3	81	83.5	112	4	US-09-839-542B-31
4	81	83.5	112	4	US-09-535-852-31
5	71	73.2	15	1	US-08-787-547-24
6	71	73.2	15	2	US-08-400-796-3
7	71	73.2	113	3	US-09-187-859-32
8	71	73.2	113	4	US-09-839-542B-32
9	71	73.2	113	4	US-09-535-852-32
10	71	73.2	1105	4	US-09-949-016-8227
11	71	73.2	1117	4	US-09-949-016-6148
12	52	53.6	906	4	US-09-417-039-11
13	49	50.5	555	2	US-08-453-702A-98
14	49	50.5	556	1	US-07-998-003A-98
15	49	50.5	556	1	US-08-453-274B-98
16	49	50.5	556	1	US-08-453-695A-98
17	49	50.5	556	1	US-08-268-161A-98
18	49	50.5	556	3	US-09-099-639-98
19	48	49.5	841	4	US-09-949-016-9669
20	45	46.4	619	4	US-09-578-441-3
21	45	46.4	620	4	US-09-578-441-4
22	45	46.4	884	2	US-08-474-067-8
23	45	46.4	884	2	US-08-474-068A-8
24	45	46.4	884	2	US-08-472-481-7
25	43	44.3	913	2	US-08-474-067-6
26	43	44.3	913	2	US-08-474-068A-6
27	43	44.3	913	2	US-08-472-481-5

```

28 42 43.3 840 4 US-09-949-016-6569 Sequence 6569, Ap
29 42 43.3 878 1 US-08-237-913-2 Sequence 2, Appli
30 42 43.3 878 3 US-08-732-429-2 Sequence 2, Appli
31 42 43.3 878 4 US-09-798-267-2 Sequence 3, Appli
32 42 43.3 878 4 US-09-798-267-3 Sequence 2, Appli
33 42 43.3 878 5 PCT-US95-05518-2 Sequence 2, Appli
34 42 43.3 916 4 US-09-949-016-8291 Sequence 8291, Ap
35 42 43.3 916 4 US-09-949-016-8292 Sequence 8292, Ap
36 41 42.3 163 4 US-09-270-767-35623 Sequence 35623, A
37 41 42.3 163 4 US-09-270-767-50840 Sequence 50840, A
38 41 42.3 208 3 US-09-134-001C-5068 Sequence 5068, Ap
39 41 42.3 301 4 US-09-710-279-1826 Sequence 1826, Ap
40 41 42.3 626 1 US-07-938-782A-2 Sequence 2, Appli
41 41 42.3 626 1 US-08-630-524-2 Sequence 2, Appli
42 41 42.3 626 4 US-09-578-441-5 Sequence 5, Appli
43 41 42.3 626 5 PCT-US93-08131-2 Sequence 2, Appli
44 41 42.3 629 4 US-09-578-441-2 Sequence 2, Appli
45 41 42.3 629 4 US-09-799-875-17 Sequence 17, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-949-016-8229
; Sequence 8229, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8229
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8229

```

```

Query Match 100.0%; Score 97; DB 4; Length 1014;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches: 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 EPNHLSKIAFKIVSQEPA 19
Db 201 EPNHLSKIAFKIVSQEPA 219

```

```

RESULT 2
US-09-187-859-31
; Sequence 31, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 112

```

Tue Mar 15 12:10:22 2005

us-10-799-005a-1.open.ra1

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-187-859-31
Query Match      83.5%; Score 81; DB 3; Length 112;
Best Local Similarity 78.9%; Pred. No. 8.9e-07;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 EPNHLNSKIAFKIVSQEPA 19
      |||:|||||:|||||:|||||:
Db      29 EPNNLSKIAFKIIRQEPS 47

RESULT 3
US-09-839-542B-31
; Sequence 31, Application US/09839542B
; Patent No. 6569996
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 31
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-839-542B-31
Query Match      83.5%; Score 81; DB 4; Length 112;
Best Local Similarity 78.9%; Pred. No. 8.9e-07;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 EPNHLNSKIAFKIVSQEPA 19
      |||:|||||:|||||:|||||:
Db      29 EPNNLSKIAFKIIRQEPS 47

RESULT 4
US-09-535-852-31
; Sequence 31, Application US/09535852
; Patent No. 6638911
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James M.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C6
; CURRENT APPLICATION NUMBER: US/09/535,852
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-535-852-31
Query Match      83.5%; Score 81; DB 4; Length 112;
Best Local Similarity 78.9%; Pred. No. 8.9e-07;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 EPNHLNSKIAFKIVSQEPA 19
      |||:|||||:|||||:|||||:
Db      29 EPNNLSKIAFKIIRQEPS 47

RESULT 5
US-08-787-547-24
; Sequence 24, Application US/08787547
; Patent No. 5783567
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; APPLICANT: Curley, Joanne M.
; APPLICANT: Langer, Robert S.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY
; TITLE OF INVENTION: OF NUCLEIC ACID
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,547
; FILING DATE: 22-JAN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 08191/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-787-547-24
Query Match      73.2%; Score 71; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 LNSKIAFKIVSQEPA 19
      |||:|||||:|||||:|||||:
Db      1 LNSKIAFKIVSQEPA 15

RESULT 6
US-08-400-796-3
; Sequence 3, Application US/08400796
; Patent No. 5874531
; GENERAL INFORMATION:
; APPLICANT: STROMINGER, JACK L.
; APPLICANT: WUCHERPFENNIG, KAI
; TITLE OF INVENTION: IDENTIFICATION OF SELF AND NON-SELF
; TITLE OF INVENTION: ANTIGENS IMPLICATED IN AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:

```

```
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/400,796
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: TWOMEY, MICHAEL J.
/ REGISTRATION NUMBER: 38,349
/ REFERENCE/DOCKET NUMBER: H0498/7015
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-720-3500
/ TELEFAX: 617-720-2441
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 15 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ FRAGMENT TYPE: internal
/ ORIGINAL SOURCE:
/ ORGANISM: HOMO SAPIENS
/
US-08-400-796-3

Query Match 73.2%; Score 71; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
Db 1 LNSKIAFKIVSQEPA 15

RESULT 7
US-09-187-859-32
/ Sequence 32, Application US/09187859A
/ Patent No. 6358920
/ GENERAL INFORMATION:
/ APPLICANT: Blaschuk, Orest W.
/ APPLICANT: Gour, Barbara J.
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
/ TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
/ FILE REFERENCE: 100086.407C1
/ CURRENT APPLICATION NUMBER: US/09/187,859A
/ CURRENT FILING DATE: 1998-11-06
/ NUMBER OF SEQ ID NOS: 4052
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 32
/ LENGTH: 113
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/
US-09-187-859-32

Query Match 73.2%; Score 71; DB 3; Length 113;
Best Local Similarity 73.7%; Pred. No. 5.3e-05;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
Db 29 EPNLNSKISYRIVSLEPA 47

RESULT 8
US-09-839-542B-32
/ Sequence 32, Application US/09839542B
/ Patent No. 6569996
/ GENERAL INFORMATION:
/ APPLICANT: Blaschuk, Orest W.
/ APPLICANT: Symonds, James Matthew
/ APPLICANT: Gour, Barbara J.
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
/ TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
/ FILE REFERENCE: 100086.407D1
/ CURRENT APPLICATION NUMBER: US/09/839,542B
/ CURRENT FILING DATE: 2001-04-20
/ NUMBER OF SEQ ID NOS: 4052
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 32
/ LENGTH: 113
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/
US-09-839-542B-32

Query Match 73.2%; Score 71; DB 4; Length 113;
Best Local Similarity 73.7%; Pred. No. 5.3e-05;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
Db 29 EPNLNSKISYRIVSLEPA 47

RESULT 9
US-09-535-852-32
/ Sequence 32, Application US/09535852
/ Patent No. 6638911
/ GENERAL INFORMATION:
/ APPLICANT: Blaschuk, Orest W.
/ APPLICANT: Symonds, James M.
/ APPLICANT: Gour, Barbara J.
/ TITLE OF INVENTION: COMPOUNDS AND ENTHODS FOR MODULATING
/ TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
/ FILE REFERENCE: 100086.407C6
/ CURRENT APPLICATION NUMBER: US/09/535,852
/ CURRENT FILING DATE: 2001-05-21
/ NUMBER OF SEQ ID NOS: 2009
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 32
/ LENGTH: 113
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/
US-09-535-852-32

Query Match 73.2%; Score 71; DB 4; Length 113;
Best Local Similarity 73.7%; Pred. No. 5.3e-05;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
Db 29 EPNLNSKISYRIVSLEPA 47

RESULT 10
US-09-949-016-8227
/ Sequence 8227, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 8227
/ LENGTH: 1105
```

Query Match 53.6%; Score 52; DB 4; Length 906;
Best Local Similarity 47.4%; Pred. No. 1.4;
Matches 9; Conservative 5; Mismatches 5; Indels

RESULT 14
US-07-998-003A-98
; Sequence 98, Application US/07998003A
; Patent No. 5643781
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESSEE: Bicknell
; STREET: 20 South Clark Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/998,003A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 564378land, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 30903
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/346-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-998-003A-98

Query Match 50.5%; Score 49; DB 1; Length 556;
Best Local Similarity 42.1%; Pred. No. 2.7;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
Db 137 DPNALNGMLRYRILSQAPS 155

RESULT 15

US-08-453-274B-98
Sequence 98, Application US/08453274B
Patent No. 5663300
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,274B
FILING DATE: 30-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: No. 5663300and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32660
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-274B-98

Query Match 50.5%; Score 49; DB 1; Length 556;
Best Local Similarity 42.1%; Pred. No. 2.7;

Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 EPNHLSKIAFKIVSQEPA 19
Db 137 DPNALNGMLRYRILSQAPS 155
Search completed: March 15, 2005, 11:55:30
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 12:04:02 ; Search time 38 Seconds
(without alignments)
48.108 Million cell updates/sec

Title: US-10-799-005A-1
Perfect score: 97
Sequence: 1 EPNHLNSKIAPKIVSQEPA 19
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 3436

Minimum DB seq length: 0
Maximum DB seq length: 19

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	26.8	17	2	PH1822
2	25	25.8	10	2	S66458
3	25	25.8	15	2	S43321
4	23	23.7	14	2	A49018
5	23	23.7	17	2	PC3196
6	22	22.7	9	2	JP0073
7	22	22.7	10	2	F44644
8	22	22.7	13	2	E39778
9	22	22.7	13	2	G44644
10	22	22.7	16	2	CQR1
11	22	22.7	16	2	PL0124
12	21.5	22.2	19	2	A37968
13	21	21.6	7	2	A15398
14	20	20.6	13	2	JQ2309
15	20	20.6	13	2	JQ2319
16	20	20.6	14	2	PH0795
17	20	20.6	14	2	PH0776
18	20	20.6	15	2	PA0097
19	20	20.6	16	2	PH0777
20	20	20.6	18	2	C40433
21	20	20.6	18	2	S70611
22	19.5	20.1	17	2	C43599
23	19	19.6	12	2	S71034
24	19	19.6	12	2	F84132
25	19	19.6	13	2	S57571
26	19	19.6	13	2	E42762
27	19	19.6	13	2	PC1008
28	19	19.6	13	2	A86126
29	19	19.6	15	2	PS0452

30 19 19.6 15 2 A36212 carboxypeptidase B
31 19 19.6 19 2 C40634 orf21 5' of eryK -
32 18 18.6 9 2 S70345 amine oxidase (cop
33 18 18.6 10 2 PT0309 IG heavy chain CRD
34 18 18.6 10 2 A24407 amicyanin - Paraco
35 18 18.6 11 2 E50691 phycobilisome 8K 1
36 18 18.6 12 2 D28551 hypothetical prote
37 18 18.6 12 2 A35585 cytochrome-binding
38 18 18.6 12 2 S27024 Na+/K+-exchanging
39 18 18.6 14 2 B29743 translation initia
40 18 18.6 14 2 PH0801 T-cell receptor al
41 18 18.6 15 2 PQ0174 stylar glycoprotei
42 18 18.6 15 2 PS0251 15K protein S106 -
43 18 18.6 15 2 A36315 recycling receptor
44 18 18.6 15 2 PH0760 T-cell receptor be
45 18 18.6 15 2 PA0054 protein QP200017 -

ALIGNMENTS

RESULT 1

PH1822
T cell receptor alpha chain V region (clone 5PBL V alpha 24-5) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: PH1822
J:Porcelli, S.; Vockey, C.E.; Brenner, M.B.; Balk, S.P.
J:Exp. Med. 178, 1-16, 1993
A:Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood

A:Reference number: PH1754; MUID:93301585; PMID:8391057
A:Accession: PH1822
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-17 <P>
Query Match 26.8%; Score 26; DB 2; Length 17;
Best Local Similarity 60.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 PNHLSKIAP 11

DB 7 PNSSASKIIF 16

RESULT 2

S66458
ferredoxin - Rhizobium meliloti (fragment)
C:Species: Rhizobium meliloti
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S66458
R:Riedel, K.U.; Jouanneau, Y.; Masepohl, B.; Puehler, A.; Klipp, W.
Eur. J. Biochem. 231, 742-746, 1995
A:Title: A Rhizobium meliloti ferredoxin (FdXN) purified from Escherichia coli donates

A:Reference number: S66458; MUID:95377307; PMID:7649175
A:Accession: S66458
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <RIE>
C:Genetics:
A:Gene: fdxN
Query Match 25.8%; Score 25; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 AFKIVSQE 17

DB 1 AFKIIASQ 8

RESULT 3

S43321
RNA-binding protein - garden pea (fragment)
C:Species: Pisum sativum (garden pea)
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 09-Jul-2004
C:Accession: S43321
R:Lakhani, S.; Khanna, N.C.; Tewari, K.K.
Plant Mol. Biol. 23, 963-979, 1993
A:Title: Nascent transcript-binding protein of the pea chloroplast transcriptionally active
A:Reference number: S43321; MUID:94083566; PMID:8260634
A:Accession: S43321
A:Molecule type: protein
A:Residues: 1-15 <LAK>
A:Cross-references: UNIPROT:Q9T2J9
C:Function:
C:Description: probably involved in the transcription of chloroplast genes
C:Keywords: RNA binding

Query Match 25.8%; Score 25; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 6.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
Db 2 NHINGTI 8

RESULT 4
A49018
myosin heavy chain, fast skeletal muscle type X - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 13-Aug-1999
C:Accession: A49018; S32161
R:DeNardi, C.; Ausoni, S.; Moretti, P.; Gorza, L.; Velleca, M.; Schiaffini
J. Cell Biol. 123, 823-835, 1993
A:Title: Type 2X-myosin heavy chain is coded by a muscle fiber type-specific and develop
A:Reference number: A49018; MUID:94043465; PMID:8227143
A:Accession: A49018
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-14 <DEN>
A:Cross-references: GB:X72591; NID:g288645; PIDN:CAA51189.1; PID:g288646
R:DeNardi, C.; Ausoni, S.; Moretti, P.; Gorza, L.; Velleca, M.; Merlie, J.; Buckingham,
submitted to the EMBL Data Library, March 1993
A:Description: Type 2X myosin heavy chain is coded by a muscle fiber type-specific and d
A:Reference number: S32161
A:Accession: S32161
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-14 <DEN2>
A:Cross-references: EMBL:X72591; NID:g288645; PIDN:CAA51189.1; PID:g288646
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: skeletal muscle

Query Match 23.7%; Score 23; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 12 KIVSQE 17
Db 9 KIISSE 14

RESULT 5
PC2196
zymogen granule membrane associated protein, ZAP47 - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: PC2196
R:Fukuda, S.
Biochim. Biotechnol. Biochem. 58, 1282-1285, 1994
A:Title: Analysis of ZAPs, zymogen granule membrane associated proteins, in the regulate
A:Reference number: PC2195; MUID:94362286; PMID:7765250
A:Accession: PC2196

A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-17 <FUK>

Query Match 23.7%; Score 23; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EPNHLN 6
Db 5 DENHIN 10

RESULT 6

JP0073
ribosomal protein L32 - Leuconostoc mesenteroides (fragment)

C:Species: Leuconostoc mesenteroides
C:Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: JP0073
R:Ochi, K.

submitted to JIPID, February 1994
A:Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal p
A:Reference number: JP0042
A:Accession: JP0073
A:Molecule type: protein
A:Residues: 1-9 <OCH>
A:Cross-references: UNIPROT:Q7M117
C:Keywords: protein biosynthesis; ribosome

Query Match 22.7%; Score 22; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNHLNSK 8
Db 3 PSNKNKSK 9

RESULT 7

F44644

neurotoxin-associated protein type B Hn+ 35K chain, band 3a - Clostridium botulinum (fra
C:Species: Clostridium botulinum
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C:Accession: F44644

R:Somers, E.; DasGupta, B.R.

J. Protein Chem. 10, 415-425, 1991
A:Title: Clostridium botulinum types A, B, Cl, and E produce proteins with or without he
A:Reference number: A44644; MUID:92143938; PMID:1781887
A:Contents: type B
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOM>
A:Cross-references: UNIPROT:Q9R5N6
A:Note: sequence extracted from NCBI backbone (NCBIP:83787)
C:Keywords: hemagglutinin

Query Match 22.7%; Score 22; DB 2; Length 10;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
Db 3 NILNDKI 9

RESULT 8

E39778

lactose phosphotransferase system lacR protein - Lactococcus lactis (fragment)
C:Species: Lactococcus lactis

C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 09-Jul-2004

C:Accession: E39778

R:van Rooijen, R.J.; van Schalkwijk, S.; de Vos, W.M.

J. Biol. Chem. 266, 7176-7181, 1991
A:Title: Molecular cloning, characterization, and nucleotide sequence of the tagatose 6-phosphate isomerase from *Rhizobium meliloti*
A:Reference number: A39778; MUID:91201377; PMID:1901863
A:Accession: E39778
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-13 <VAN>
A:Cross-references: UNIPROT:P18816; GB:J05748
C:Superfamily: regulatory protein gntR
C:Keywords: DNA binding; transcription regulation

Query Match 22.7%; Score 22; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EPNHLSK 8
| | | | |
Db 3 ESLHWKK 10

RESULT 9
G44644
neurotoxin-associated protein type B Hn+ 35K chain, band 3b - Clostridium botulinum (fra
C:Species: Clostridium botulinum
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
A:Accession: G44644
R:Somers, E.; DasGupta, B.R.
J. Protein Chem. 10, 415-425, 1991
A:Title: Clostridium botulinum types A, B, C1, and E produce proteins with or without he
A:Reference number: A44644; MUID:92143938; PMID:1781887
A:Contents: type B
A:Accession: G44644
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-13 <SOM>
A:Cross-references: UNIPROT:Q9RSN5
A:Note: sequence extracted from NCBI backbone (NCBIP:83785)
C:Keywords: hemagglutinin

Query Match 22.7%; Score 22; DB 2; Length 13;
Best Local Similarity 71.4%; Pred. No. 1.8e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NHLNSKI 9
| | | | |
Db 6 NILNDKI 12

RESULT 10
CORT
cerebellin - rat
N:Contains: des-Ser-cerebellin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 09-Jul-2004
A:Accession: A03135
R:Slemmon, J.R.; Blacher, R.; Danho, W.; Hempstead, J.L.; Morgan, J.I.
Proc. Natl. Acad. Sci. U.S.A. 81, 6866-6870, 1984
A:Title: Isolation and sequencing of two cerebellum-specific peptides.
A:Reference number: A03135
A:Accession: A03135
A:Molecule type: protein
A:Residues: 1-16 <SLE>
A:Cross-references: UNIPROT:P23436
C:Comment: The sequence of des-Ser-cerebellin lacks 1-Ser.
C:Comment: The cerebellin is localized in the Purkinje cells of rat cerebellum.
C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
C:Keywords: brain

Query Match 22.7%; Score 22; DB 2; Length 16;
Best Local Similarity 33.3%; Pred. No. 2.3e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 NSKIAFKIV 14

Db 3 SAKVAFSAI 11
| | | | |
RESULT 11
PL0124
cerebellin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Aug-2004
A:Accession: PL0124
R:Yiangou, Y.; Burnet, P.; Nikou, G.; Chrysanthou, B.J.; Bloom, S.R.
J. Neurochem. 53, 886-889, 1989
A:Title: Purification and characterization of cerebellins from human and porcine cerebel
A:Reference number: PL0124; MUID:89341798; PMID:2760624
A:Accession: PL0124
A:Molecule type: protein
A:Residues: 1-16 <YIA>
A:Cross-references: UNIPROT:P23435; UNIPROT:P23436
A:Experimental source: brain
A:Note: the sequences of human and porcine cerebellin are identical to that of the rat
C:Comment: The peptide is identified in two molecular forms having 16 and 15 amino acid

Query Match 22.7%; Score 22; DB 2; Length 16;
Best Local Similarity 33.3%; Pred. No. 2.3e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 NSKIAFKIV 14
| | | | |
Db 3 SAKVAFSAI 11

RESULT 12
A37968
neural surface protein Bravo - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 15-Aug-1997
A:Accession: A37968; A36345
R:de la Rosa, E.J.; Kayyem, J.F.; Roman, J.M.; Stierhof, Y.D.; Dreyer, W.J.; Schwarz, U
J. Cell Biol. 112, 1049, 1991
A:Reference number: A37968; MUID:91154309; PMID:1999455
A:Contents: erratum
A:Accession: A37968
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-19
R:de la Rosa, E.J.; Kayyem, J.F.; Roman, J.M.; Stierhof, Y.D.; Dreyer, W.J.; Schwarz, U
J. Cell Biol. 111, 3087-3096, 1990
A:Title: Topologically restricted appearance in the developing chick retinotectal system
A:Reference number: A36345; MUID:91100421; PMID:2269667
A:Accession: A36345
A:Molecule type: protein
A:Residues: 1-7,9-19 <DE2>

Query Match 22.2%; Score 21.5; DB 2; Length 19;
Best Local Similarity 42.9%; Pred. No. 3.3e+03;
Matches 6; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 5 LNSKIAPKIVSQEP 18
| | | | |
Db 5 LDSKLX-EXLSQPP 17

RESULT 13
A15398
choline oxidase (EC 1.1.3.17) - Alcaligenes sp. (tentative sequence) (fragment)
C:Species: Alcaligenes sp.
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
A:Accession: A15398
R:Ohira-Fukuyama, M.; Miyake, Y.; Emi, S.; Yamano, T.
J. Biochem. 88, 197-203, 1980
A:Title: Identification and properties of the prosthetic group of choline oxidase from
A:Reference number: A15398; MUID:81006769; PMID:6997283
A:Accession: A15398

Search completed: March 15, 2005, 12:13:20
Job time : 40 secs

A:Molecule type: protein
A:Residues: 1-7 <OHT>
A:Cross-references: UNIPROT:P16101
C:Keywords: oxidoreductase

Query Match 21.6%; Score 21; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNH 4
|
|
|
Db 3 PNH 5

RESULT 14
JQ2309
hypothetical 1.6K protein - tomato chloroplast (strain Toko)
C:Species: chloroplast Lycopersicon esculentum (tomato)
C:Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: JQ2309
R:Kawagoe, Y.; Kikuta, Y.
Theor. Appl. Genet. 81, 13-20, 1991
A:Title: Chloroplast DNA evolution in potato (Solanum tuberosum L.).
A:Reference number: JQ2306
A:Accession: JQ2309
A:Molecule type: DNA
A:Residues: 1-13 <RAW>
A:Cross-references: UNIPROT:Q7M2F1
A:Experimental source: strain Toko
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 20.6%; Score 20; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 4e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLN 6
Db 3 EKTHIN 8

RESULT 15
JQ2319
hypothetical 1.6K protein - potato chloroplast
C:Species: chloroplast Solanum tuberosum (potato)
C:Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: JQ2319; JQ2314
R:Kawagoe, Y.; Kikuta, Y.
Theor. Appl. Genet. 81, 13-20, 1991
A:Title: Chloroplast DNA evolution in potato (Solanum tuberosum L.).
A:Reference number: JQ2306
A:Accession: JQ2319
A:Molecule type: DNA
A:Residues: 1-13 <KW1>
A:Cross-references: UNIPROT:Q7M2E5
A:Experimental source: cv. W553-4
A:Accession: JQ2314
A:Molecule type: DNA
A:Residues: 1-13 <KW2>
A:Experimental source: cv. 150
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 20.6%; Score 20; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 4e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLN 6
Db 3 EKTHIN 8

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 11:55:39 ; Search time 168 Seconds
(without alignments)
57.914 Million cell updates/sec

Title: US-10-799-005A-1
Perfect score: 97
Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 10410

Minimum DB seq length: 0
Maximum DB seq length: 19

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03: *
1: uniprot_spot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	26.8	15	2 Q6GWV7	Q6GWV7 bos taurus
2	26	26.8	15	2 Q8C9M9	Q8C9M9 mus musculus
3	26	26.8	19	2 Q9BX20	Q9BX20 homo sapien
4	25	25.8	15	2 Q679L5	Q679L5 homo sapien
5	25	25.8	15	2 Q9UMY6	Q9UMY6 homo sapien
6	25	25.8	15	2 Q9T2J9	Q9T2J9 pisum sativ
7	25	25.8	16	2 Q8WXZ7	Q8WXZ7 homo sapien
8	25	25.8	16	2 Q7RPN7	Q7RPN7 plasmodium
9	25	25.8	18	2 Q43834	Q43834 homo sapien
10	24	24.7	9	2 Q6EX64	Q6EX64 hyptis flor
11	24	24.7	10	2 Q9RSN4	Q9RSN4 clostridium
12	24	24.7	13	1 PROX_ORYSA	P83647 oryza sativ
13	24	24.7	16	1 FOR1_MYRGU	P81438 myrmecia gu
14	24	24.7	16	1 FOR2_MYRGU	P81437 myrmecia gu
15	24	24.7	19	2 Q36925	Q36925 nicotiana p
16	24	24.7	19	2 Q7GEF1	Q7GEF1 nicotiana v
17	24	24.7	19	2 Q7GEF2	Q7GEF2 nicotiana v
18	24	24.7	19	2 Q9G466	Q9G466 brassica ju
19	23	23.7	12	2 Q8MJEA	Q8MJEA aotus azara
20	23	23.7	13	2 Q957T7	Q957T7 trichopsis
21	23	23.7	14	2 Q06415	Q06415 rattus ratt
22	23	23.7	15	2 Q8ST73	Q8ST73 drosophila
23	23	23.7	15	2 Q9TR17	Q9TR17 oryctolagus
24	23	23.7	16	2 Q7S0C4	Q7S0C4 neurospora
25	23	23.7	17	2 Q9T2S0	Q9T2S0 solanum tub
26	23	23.7	18	2 Q9TR05	Q9TR05 bos taurus
27	23	23.7	18	2 Q6VFR3	Q6VFR3 squirrel po
28	23	23.7	19	2 Q9S8Q0	Q9S8Q0 pinus sylve
29	22	22.7	9	2 Q7OYAT	Q7OYAT alvestia ros
30	22	22.7	9	2 Q7M117	Q7M117 leuconostoc
31	22	22.7	10	2 Q9R5N6	Q9R5N6 clostridium

32	22	22.7	11	2 P89269	P89269 xestia c-ni
33	22	22.7	13	2 Q9AIR1	Q9AIR1 pseudomonas
34	22	22.7	13	2 Q9RSN5	Q9RSN5 clostridium
35	22	22.7	14	1 UC15_MAIZE	P80621 zea mays (m
36	22	22.7	14	2 Q06414	Q06414 rattus ratt
37	22	22.7	14	2 Q9QVF3	Q9QVF3 rattus sp;
38	22	22.7	15	1 THBI_BOTJA	P84026 bothrops ja
39	22	22.7	15	2 Q9QVG2	Q9QVG2 mus sp. dif
40	22	22.7	16	1 CERB_PIG	P63181 sus scrofa
41	22	22.7	16	1 CERB_RAT	P63182 rattus norv
42	22	22.7	16	2 Q9TZR2	Q9TZR2 haliotis ru
43	22	22.7	17	2 Q8N6B0	Q8N6B0 homo sapien
44	22	22.7	17	2 P83329	P83329 streptococc
45	22	22.7	17	2 Q7X3X0	Q7X3X0 streptomyce

ALIGNMENTS

RESULT 1

Q6GWV7 PRELIMINARY; PRT; 15 AA.
AC Q6GWV7; (Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE Interleukin 8 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Meade K.G., Fitzgerald D.C., Murphy E.P., Baird A.W., Machugh D.E.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY627308; AAT47550.1; -.
FT NON_TER 15
SQ SEQUENCE 15 AA; 1504 MW; DA1932E93C487C14 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 15;
Best Local Similarity 33.3%; Pred. No. 3e+03; Mismatches 5; Indels 0; Gaps 0;
Matches 5; Conservative 5;

Qy 5 LNSKIAFKIVSQEPA 19

Db 1 MTSKLAVALLAAFP 15

RESULT 2

Q8C9M9 PRELIMINARY; PRT; 15 AA.
ID Q8C9M9
AC Q8C9M9; (Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone: A630035G04 product: DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 26, full insert sequence. (Fragment).
DE Name=DDx26;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Thymus;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.

```
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Satoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK041758; BAC31055.1; -
DR MGB; MGI:202397; Ddx26.
DR GO; GO:0005615; C:extracellular space; TAS.
FT NON TER 1
SQ SEQUENCE 15 AA; 1816 MW; CDF3F69D9D63857 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NHLNS 7
DB 10 NHINS 14

RESULT 3
Q9BX20 PRELIMINARY; PRT; 19 AA.
ID Q9BX20
AC Q9BX20
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Satoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK041758; BAC31055.1; -
DR MGB; MGI:202397; Ddx26.
DR GO; GO:0005615; C:extracellular space; TAS.
FT NON TER 1
SQ SEQUENCE 15 AA; 1816 MW; CDF3F69D9D63857 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NHLNS 7
DB 10 NHINS 14

RESULT 3
Q9BX20 PRELIMINARY; PRT; 19 AA.
ID Q9BX20
AC Q9BX20
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
```

```
DE BAA70C13.1 (KIAA1272 similar to rat tulip proteins 1 and 2)
DE (Fragment).
GN Name=BAA70C13.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Peck A.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161658; CAC36066.1; -
FT NON TER 1
SQ SEQUENCE 19 AA; 2211 MW; 2BB8DB0ABFF5A5C0 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 19;
Best Local Similarity 36.4%; Pred. No. 3.8e+03;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 8 KIAFKIVSOEP 18
DB 5 KVSFYLSKGP 15

RESULT 4
Q679L5 PRELIMINARY; PRT; 15 AA.
ID Q679L5
AC Q679L5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE ABO glycosyltransferase (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hosseini-Maaf B., Hellberg A., Rodrigues M.J., Chester A.M.,
RA Olsson M.L.;
RT "ABO Exon and Intron Analysis in Individuals with the ABO Phenotype
Reveals a Novel O1v-A2 Hybrid Allele that Causes Four Missense
Mutations in the A Transferase.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY377125; AAR24575.1; -
DR GO; GO:0016740; F:transferase activity; IEA.
FT NON TER 1
FT NON TER 15
SQ SEQUENCE 15 AA; 1729 MW; FB7670F419288D18 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 4.4e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHL 5
DB 5 EPDHL 9

RESULT 5
Q9UMY6 PRELIMINARY; PRT; 15 AA.
ID Q9UMY6
AC Q9UMY6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ABO glycosyltransferase (Fragment).
GN Name=ABO;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

RN  SEQUENCE FROM N.A.
RX  MEDLINE=97318856; PubMed=9175793; DOI=10.1006/bbrc.1997.6713;
RA  Olsson M.L., Guerreiro J.F., Zago M.A., Chester M.A.;
RT  "Molecular analysis of the O alleles at the blood group ABO locus in
FT  populations of different ethnic origin reveals novel crossing-over
RT  events and point mutations.";
RL  Biochem. Biophys. Res. Commun. 234:779-782(1997).
RN  SEQUENCE FROM N.A.
RP  Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF170890; AAD51650.1; -.
DR  GO; GO:00016740; F:transferase activity; IEA.
KW  Transferase.
FT  NON_TER 1 15
FT  NON_TER 15 15
SQ  SEQUENCE 15 AA; 1762 MW; FB7670E69CC88D18 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 4.4e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHL 5
Db ||:|
5 EPDHL 9

RESULT 6
Q9T2J9 PRELIMINARY; PRT; 15 AA.
AC Q9T2J9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transcript-binding protein (Fragment).
OS Pisum sativum (Garden pea).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN  SEQUENCE
RP MEDLINE=94083566; PubMed=8260634;
RA Lakhani S., Khanna N.C., Tewari K.K.;
RT "Nascent transcript-binding protein of the pea chloroplast
RT transcriptionally active chromosome.";
RL Plant Mol. Biol. 23:963-979(1993).
DR PIR; S43321; S43321.
SQ SEQUENCE 15 AA; 1668 MW; 0A778A7774F79D79 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 4.4e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NHLSKI 9
Db ||:|
2 NHNGTI 8

RESULT 7
Q9WXZ7 PRELIMINARY; PRT; 16 AA.
AC Q9WXZ7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha-3-galactosylaminyltransferase (EC 2.4.1.40) (Fragment).
GN Name=ABO;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN  SEQUENCE FROM N.A.
RP MEDLINE=21411478; PubMed=11520811; DOI=10.1182/blood.V98.5.1585;
RA Olsson M.L., Irehaid N.M., Hosseini-Maaf B., Hellberg A., Moulds M.K.,
RA Sareneva H., Chester M.A.;
RT "Genomic analysis of clinical samples with serologic ABO blood
RT grouping discrepancies: identification of 15 novel A and B subgroup
RT alleles.";
RL Blood 98:1585-1593(2001).
DR EMBL; AF324008; AAL37340.1; -.
DR GO; GO:0004380; F:glycoprotein-fucosylgalactoside alpha-N-ace. .; IEA.
DR GO; GO:0016757; F:transferase activity; transferring glycosyl. .; IEA.
KW Glycosyltransferase; Transferase.
FT NON_TER 1 16
FT NON_TER 16 16
SQ SEQUENCE 16 AA; 1849 MW; 67BB7670F40CC88D CRC64;

Query Match 25.8%; Score 25; DB 2; Length 16;
Best Local Similarity 80.0%; Pred. No. 4.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHL 5
Db ||:|
5 EPDHL 9

RESULT 8
Q7RPN7 PRELIMINARY; PRT; 16 AA.
AC Q7RPN7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY01420;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN  SEQUENCE FROM N.A.
RP STRAIN=17XNL;
RC PubMed=12368865; DOI=10.1038/nature01099;
RX Carlton J.M., Anguoli S.V., Suh B.B., Kooij T.W., Perteu M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabai A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000375; EAA20751.1; -.
KW Hypothetical protein.
SQ SEQUENCE 16 AA; 1784 MW; E59C6B328ED05081 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 16;
Best Local Similarity 40.0%; Pred. No. 4.7e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
Db :||:|
4 VKIEVSAEP 13

RESULT 9

```

```
043834
ID O43834 PRELIMINARY; PRT; 18 AA.
AC O43834;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Heparan sulfate proteoglycan (Fragment).
GN Name=HSPG2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Glomerular basement membrane;
RX MEDLINE=97458991; PubMed=9313766;
RA Hansen P.M., Chowdhury T., Deckert T., Helligren A., Bain S.C.,
RA Pociot F.;
RT "Genetic variation of the heparan sulfate proteoglycan gene (perlecan
RT gene). Association with urinary albumin excretion in IDDM patients.";
RL Diabetes 46:1658-1659(1997).
DR EMBL; U41286; AAB9516.1; -.
FT NON TER 1
FT NON TER 18
SQ SEQUENCE 18 AA; 1869 MW; FBFFBA7D8232263E CRC64;

Query Match 25.8%; Score 25; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.4e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 SOEPA 19
Db 4 SOEPA 8

RESULT 10
Q6EX64 PRELIMINARY; PRT; 9 AA.
ID Q6EX64;
AC Q6EX64;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Ribosomal protein (Fragment).
GN Name=rps16;
OS Hyptis floribunda.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Hyptis.
OX NCBI_TaxID=204126;
RN [1]
RP SEQUENCE FROM N.A.
RA Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Willis F.,
RA Powell M.P., Savolainen V.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ505339; CAD45462.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Chloroplast; Ribosomal protein.
FT NON TER 1
FT NON TER 9
SQ SEQUENCE 9 AA; 952 MW; 45A858786415B447 CRC64;

Query Match 24.7%; Score 24; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLNSK 8
Db 1 HLNSR 5

RESULT 11
Q9RSN4 PRELIMINARY; PRT; 10 AA.
ID Q9RSN4;
AC Q9RSN4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Neurotoxin type A HN+ 35 kDa subunit (Fragment).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE.
RX MEDLINE=92143938; PubMed=1781887;
RA Somers E., DasGupta B.R.;
RT "Clostridium botulinum types A, B, C1, and E produce proteins with or
RT without hemagglutinating activity: do they share common amino acid
RT sequences and genes?";
RL J. Protein Chem. 10:415-425(1991);
FT NON TER 1
FT NON TER 10
SQ SEQUENCE 10 AA; 1143 MW; CF3CB4A44735B456 CRC64;

Query Match 24.7%; Score 24; DB 2; Length 10;
Best Local Similarity 71.4%; Pred. No. 4.3e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
Db 4 NSLNDKI 10

RESULT 12
PROX_ORYSA STANDARD; PRT; 13 AA.
ID PROX_ORYSA;
AC P83647;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Probable profilin LP04 (Fragments).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Indica / IR64; TISSUE=Panicle;
RA Hosseini Salekdeh S.G., Bennett J.;
RT "Proteome analysis of rice panicle.";
RL Submitted (JUL-2003) to Swiss-Prot.
CC -!- FUNCTION: Binds to actin and affects the structure of the
CC cytoskeleton. At high concentrations, profilin prevents the
CC polymerization of actin, whereas it enhances it at low
CC concentrations. By binding to p1p2, it inhibits the formation of
CC IP3 and DG (By similarity).
CC -!- SUBUNIT: Occurs in many kinds of cells as a complex with monomeric
CC actin in a 1:1 ratio.
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 4.4, its MW is: 14.0 kDa.
CC -!- SIMILARITY: Belongs to the profilin family.
DR InterPro; IPR002097; Profilin.
DR PROSITE; PS00414; PROFILIN; PARTIAL.
KW Actin-binding; Cytoskeleton; Direct protein sequencing;
FT NON TER 1
FT NON TER 5
FT NON TER 13
SQ SEQUENCE 13 AA; 1362 MW; 0A3022EE0E52C68B CRC64;

Query Match 24.7%; Score 24; DB 1; Length 13;
Best Local Similarity 30.0%; Pred. No. 5.6e+03;
Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```

QY 9 IAFKIVSQEP 18
   :|: :|:
Db 4 LAYVIOQEP 13

RESULT 13
FOR1 MYRGU STANDARD; PRT; 16 AA.
ID P81438;
AC P81438;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Formaeicin 1.
OS Myrmecia gulosa (Red bulldog ant).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Vespoidea;
OC Formicidae; Myrmecilinae; Myrmecilini; Myrmecia.
OX NCBI_TaxID=36170;
RN [1]
RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
RC TISSUE=Hemolymph;
RX MEDLINE=98165787; PubMed=9497332; DOI=10.1074/jbc.273.11.6139;
RA Mackintosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
RT "Isolation from an ant Myrmecia gulosa of two inducible O-glycosylated
RT proline-rich antibacterial peptides."
RL J. Biol. Chem. 273:6139-6143(1998).
CC -!- FUNCTION: Antibacterial peptide. Has activity against E.coli but
CC none against other Gram-negative bacteria and Gram-positive
CC bacteria.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- INDUCTION: By bacterial infection.
CC -!- PTM: O-linked glycan consists of a Gal-GalNAc disaccharide, O-
CC glycosylation is essential for full biological activity.
KW Antibiotic; Direct protein sequencing; Glycoprotein; Hemolymph;
KW Insect immunity.
FT CARBOHYD 11 11 O-linked (GalNAc...).
SQ SEQUENCE 16 AA; 1794 MW; 80CEA3AABBC2E0AE CRC64;

Query Match 24.7%; Score 24; DB 1; Length 16;
Best Local Similarity 57.1%; Pred. No. 7e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLNSK 8
   ||:|:|
Db 3 PNPVANK 9

RESULT 14
FOR2 MYRGU STANDARD; PRT; 16 AA.
ID P81437;
AC P81437;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Formaeicin 2.
OS Myrmecia gulosa (Red bulldog ant).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Vespoidea;
OC Formicidae; Myrmecilinae; Myrmecilini; Myrmecia.
OX NCBI_TaxID=36170;
RN [1]
RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
RC TISSUE=Hemolymph;
RX MEDLINE=98165787; PubMed=9497332; DOI=10.1074/jbc.273.11.6139;
RA Mackintosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
RT "Isolation from an ant Myrmecia gulosa of two inducible O-glycosylated
RT proline-rich antibacterial peptides."
RL J. Biol. Chem. 273:6139-6143(1998).
CC -!- FUNCTION: Antibacterial peptide. Has activity against E.coli but
CC none against other Gram-negative bacteria and Gram-positive
CC bacteria.
CC -!- SUBCELLULAR LOCATION: Secreted.

```

```

CC -!- INDUCTION: By bacterial infection.
CC -!- PTM: O-linked glycan consists of a Gal-GalNAc disaccharide, O-
CC glycosylation is essential for full biological activity.
KW Antibiotic; Direct protein sequencing; Glycoprotein; Hemolymph;
KW Insect immunity.
FT CARBOHYD 11 11 O-linked (GalNAc...).
SQ SEQUENCE 16 AA; 1807 MW; 9C3CA3B00BC2E0AE CRC64;

Query Match 24.7%; Score 24; DB 1; Length 16;
Best Local Similarity 57.1%; Pred. No. 7e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLNSK 8
   ||:|:|
Db 3 PNPVANK 9

RESULT 15
Q36925 PRELIMINARY; PRT; 19 AA.
ID Q36925;
AC Q36925;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Rp19, protein (Fragment).
GN Name=rp19;
OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96397499; PubMed=8804393; DOI=10.1007/s004389670022;
RA Goulding S.E., Olmstead R.G., Morden C.W., Wolfe K.H.;
RT "Ebb and flow of the chloroplast inverted repeat."
RL Mol. Gen. Genet. 252:195-206(1996).
DR EMBL; Z71240; CAA94946.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR002222; Ribosomal_S19.
DR Pfam; PF00203; Ribosomal_S19; I.
KW Chloroplast.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2239 MW; 19AA300AB9DE1D77 CRC64;

Query Match 24.7%; Score 24; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 8.4e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NHLNSKI 9
   ||:|:|
Db 13 NHLKKKI 19

Search completed: March 15, 2005, 12:12:35
Job time : 170 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 11:54:49 ; Search time 160 Seconds
(without alignments)
45.928 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 691830

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	73.2	15	2	Aaw04843 Self epit
2	71	73.2	15	2	Aaw64815 Desmoglei
3	71	73.2	15	2	Aaw78814 Desmoglei
4	71	73.2	15	3	Aab33625 MHC class
5	71	73.2	15	4	Aag93721 Human des
6	71	73.2	15	5	Aao17033 Desmoglei
7	71	73.2	15	6	Abu96577 MHC class
8	71	73.2	15	8	Adq14318 Human des
9	71	73.2	15	8	Adr41700 Desmoglei
10	71	73.2	15	8	Adsl4310 Desmoglei
11	32	33.0	15	6	Abp57835 A. margin
12	31	32.0	15	5	Abd04739 Human hep
13	31	32.0	15	8	Adj34233 Human sec
14	31	32.0	16	8	Adj34231 Human sec
15	31	32.0	17	2	Aar21981 hIL-1 bin
16	31	32.0	17	2	Aar41906 Human IL-
17	31	32.0	17	2	Aar83473 Anti-infl
18	31	32.0	17	2	Aar83456 Immunosp
19	30.5	31.4	9	6	Abp71110 E10 prote
20	30	30.9	9	8	Adm64126 HLA bindi
21	30	30.9	9	8	Adp26187 Plasmodi
22	30	30.9	14	4	Aam97543 Human pep
23	30	30.9	15	2	Aay01783 Active pr
24	30	30.9	15	8	Adl26206 Synthetic
25	30	30.9	15	8	Adl26187 Synthetic

ALIGNMENTS

RESULT 1

AAW04843

ID AAW04843 standard; peptide; 15 AA.

AC AAW04843;

DT 18-FEB-1997 (first entry)

DE Self epitope of desmoglein 3, implicated in autoimmune disease.

KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
 KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
 KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
 KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
 KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.

OS Homo sapiens.

PN WO9627387-A1.

PD 12-SEP-1996.

PF 07-MAR-1996; 96WO-US003182.

PR 07-MAR-1995; 95US-00400796.

PA (HARD) HARVARD COLLEGE.

PI Strominger JL, Wucherpfennig KW;

XX WPI; 1996-425218/42.

DR Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
 PT - useful in disease treatment, and method for identification of other
 PT self and non-self antigens implicated in auto-immune disease.

XX Claim 1; Page 39; 58pp; English.

CC Pharmaceutical preparations for tolerisation to antigens comprise either
 CC an isolated human non-collagen or non-mysin basic protein (MBP)
 CC polypeptide which is capable of tolerising an individual to an
 CC autoantigen; or an isolated human pathogen polypeptide capable of
 CC tolerising an individual to that polypeptide. In both cases, the
 CC polypeptide (whether self or non-self) includes an amino acid sequence
 CC corresponding to a sequence motif for a MHC class II protein, such as HLA
 CC -DR, which is associated with a human autoimmune disease and which binds
 CC to the polypeptide to activate autoreactive T-cells in individuals with
 CC the autoimmune disease. This peptide is derived from the human desmoglein

Aar14398 Sclavo pe
 Aae08158 Peptide #
 Abg31019 Human fib
 Aaw85386 Helper T-
 Aaw85174 Helper T-
 Aab13871 L2/HNK1 c
 Adl26192 Synthetic
 Abp82489 G protein
 Ado34054 Human CLA
 Abr39983 Rat late
 Abo59682 Human gen
 Aae24272 Murine E-
 Abg71346 Bovine S-
 Abg71853 Bovine AK
 Aaw78190 Human sec
 Abb40564 Peptide #
 Aam34425 Peptide #
 Aam74313 Human bon
 Aam61524 Human bra
 Abg56112 Human liv

CC 3 protein (amino acids 190-204) and is implicated as a self epitope in
 CC pemphigus vulgaris. Peptides derived from the human desmoglein protein
 CC are described in AAW04841-47

XX SQ Sequence 15 AA;
 Query Match 73.2%; Score 71; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LNSKIAFKIVSQEPA 19
 |||||
 DB 1 LNSKIAFKIVSQEPA 15

RESULT 2
 AAW64815
 ID AAW64815 standard; peptide; 15 AA.

XX AC AAW64815;
 XX DT 29-SEP-1998 (first entry)
 XX DE Desmoglein-3 190-204.
 XX KW Desmoglein; DG; gene therapy; pemphigus vulgaris; microparticle;
 XX KW autoantigen; autoimmune disease; MHC.

XX OS Homo sapiens.

XX PN US5783567-A.

XX PD 21-JUL-1998.

XX PF 22-JAN-1997; 97US-00787547.

XX PR 22-JAN-1997; 97US-00787547.

XX PA (PANG-) PANGAEA PHARM INC.

XX PI Langer RS, Hedley ML, Curley JM;

XX WPI; 1998-427077/36.

XX PT Microparticle encapsulated nucleic acids - for recombinant expression of
 XX PT proteins e.g. in gene therapy.

XX PS Disclosure; Col 4; 42pp; English.

XX CC The patent describes a new preparation of microparticles each comprising
 CC a polymeric matrix and a nucleic acid. The polymeric matrix consists of
 CC one or more synthetic polymers having a solubility in water of less than
 CC 1 mg/l (e.g. poly-lactic-co-glycolic acid); and at least 90% of the
 CC microparticles have a diameter of less than 100 microns. The
 CC microparticles are useful for the delivery of nucleic acids to phagocytic
 CC cells. In one embodiment the microparticles are less than 20 microns in
 CC diameter and the nucleic acid (preferably in closed circular form)
 CC includes an expression control sequence operatively linked to a coding
 CC sequence, where the expression product of the coding sequence is a
 CC polypeptide having a length and a sequence which permits it to bind to an
 CC MHC class I or II molecule. The expression product is thus an effective
 CC stimulator of an immune response in mammals. The present sequence, an
 CC antigenic portion of desmoglein 3, is an example of an MHC class II
 CC peptide which can be expressed by the nucleic acid. It is associated with
 CC pemphigus vulgaris

XX SQ Sequence 15 AA;

Query Match 73.2%; Score 71; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LNSKIAFKIVSQEPA 19

DB 1 LNSKIAFKIVSQEPA 15
 |||||

RESULT 3
 AAW78814

ID AAW78814 standard; peptide; 15 AA.

XX AC AAW78814;

XX DT 17-NOV-1998 (first entry)

XX DE Desmoglein 3 protein fragment 190-204.

XX KW Microparticle; delivery; polymeric matrix; autoantigen; tumour antigen;
 XX KW class II associated peptide; pathogen; gene therapy; genetic disease;
 XX KW infection; downregulation; immune response.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9831398-A1.

XX PD 23-JUL-1998.

XX PF 22-JAN-1998; 98WO-US001499.

XX PR 22-JAN-1997; 97US-00787547.

XX PR 06-JAN-1998; 98US-00003253.

XX PA (PANG-) PANGAEA PHARM INC.

XX PI Hedley ML, Curley JM, Langer RS, Lunsford LB;

XX WPI; 1998-427556/36.

XX PT New preparations of microparticles - comprising a synthetic polymer
 XX PT matrix and nucleic acid comprising an expression vector for use in gene
 XX PT therapy.

XX PS Disclosure; Page 8; 101pp; English.

XX CC A microparticle preparation (MP) has been developed, consisting of
 CC microparticles having a diameter of less than 100 nm. The MP comprises:
 CC (a) a polymeric matrix (PM) consisting of one or more synthetic polymers
 CC having a solubility in water of less than 1 mg/l; and (b) an expression
 CC vector selected from RNA molecules (at least 50% of which are closed
 CC circles) or circular plasmid DNA (at least 50% of which are supercoiled).
 CC Also described is a MP of at most 20 microns in diameter, comprising: (a)
 CC a PM; and (b) a NAM comprising an expression control sequence operatively
 CC linked to a coding sequence, where the coding sequence encodes an
 CC expression product selected from: (i) a polypeptide at least 7 amino
 CC acids in length, having a sequence identical to the sequence of: (i) a
 CC fragment of a naturally-occurring mammalian protein; or (ii) a fragment
 CC of a naturally-occurring protein from an infectious agent which infects a
 CC mammal; (2) a peptide having a length and sequence which permits it to
 CC bind to an MHC class I or II molecule; and (3) the polypeptide or the
 CC peptide linked to a trafficking sequence. AAW69763 to AAW69765, and
 CC AAW7893 to AAW78897 are peptide fragments for use in the present
 CC invention. The MPs are highly effective vehicles for the delivery of
 CC polynucleotides into phagocytic cells. They can be used for gene therapy,
 CC e.g. for treating genetic diseases, infections or tumours or for
 CC downregulating an immune response

XX SQ Sequence 15 AA;

Query Match 73.2%; Score 71; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LNSKIAFKIVSQEPA 19

DB 1 LNSKIAFKIVSQEPA 15
 |||||

RESULT 4
AAB33625
ID AAB33625 standard; peptide; 15 AA.
XX
AC AAB33625;
XX
DT 26-JAN-2001 (first entry)
XX
DE MHC class II associated immunogenic peptide SEQ ID 24.
XX
KW Microparticle; nucleic acid delivery; immunogenic peptide; MHC I; MHC II;
KW major histocompatibility complex; vaginal tissue; mucosal tissue.
XX
XX Unidentified.
OS
PN WO200053161-A2.
XX
PD 14-SEP-2000.
XX
PF 10-MAR-2000; 2000WO-US006578.
XX
PR 11-MAR-1999; 99US-00266463.
PR 27-MAY-1999; 99US-00321346.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Luneford LB, Putnam D, Hedley ML;
XX
DR WPI; 2000-638130/61.
XX
PT Microparticles useful for administering a nucleic acid into the mucosal
PT tissue preferably vaginal tissue of an animal, comprises a polymeric
PT matrix, a lipid and a nucleic acid molecule.
XX
PS Claim 25; Page 11; 96pp; English.
XX
CC The present invention relates to microparticles which are less than 20
CC microns in diameter, which comprise a polymeric matrix, a lipid and a
CC nucleic acid molecule. The microparticle is specifically not encapsulated
CC in a liposome and does not comprise a cell. The nucleotide sequence
CC encodes an expression product that binds to major histocompatibility
CC complex (MHC) type I or II molecules. Peptides AAB33602-B33647 represent
CC MHC class II associated immunogenic peptides, and AAB33648-B33710
CC represent MHC class I associated immunogenic peptides. The peptides are
CC examples of the expression products of the nucleotide sequences which can
CC be included in the microparticles of the invention. Sequences AAB33711-
CC B33716 represent alternative expression products and nuclear localisation
CC signals also used in the invention. The microparticles are useful for
CC administering a nucleic acid into the mucosal tissue preferably vaginal
CC tissue of an animal
XX
SQ Sequence 15 AA;
Query Match 73.2%; Score 71; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 LNSKIAFKIVSQEPA 19
Db 1 LNSKIAFKIVSQEPA 15
RESULT 5
AAG93721
ID AAG93721 standard; peptide; 15 AA.
XX
AC AAG93721;
XX
DT 17-SEP-2001 (first entry)
XX
DE Human desmoglein 3 peptide 1.

XX Continuous flow production; microparticle; gene therapy;
KW antisense therapy; vaccination; treatment; autoimmune disease;
KW immune response modulation.
XX
OS Homo sapiens.
XX
PN WO200136583-A1.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-US031770.
XX
PR 19-NOV-1999; 99US-00443654.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Hsu Y, Tyo M;
XX
DR WPI; 2001-425203/45.
XX
PT Continuous production of microparticles containing nucleic acid for e.g.
PT gene therapy, comprises mixing a solution of polymeric material and
PT nucleic acid with a surfactant solution, removing solvent and drying.
XX
PS Disclosure; Page 9; 47pp; English.
XX
CC The present sequence is that of a peptide of the invention. The invention
CC relates to a method for scalable, continuous flow production of a nucleic
CC acid containing microparticle that maintains the structural integrity of
CC the associated nucleic acid and results in a microparticle having purity
CC suitable for introduction into an animal host. Microparticles prepared
CC according to the method can be used for delivery of a nucleic acid for
CC gene therapy, antisense therapy, vaccination, treatment of autoimmune
CC disease and either specific or non-specific modulation of an immune
CC response. The microparticles may also be used to deliver nucleic acid
CC encoding a protein or peptide useful in any kind of therapy. The method
CC is economical, aseptic and scalable. The method also enables control over
CC the size of microparticles. The microparticles produced are free of
CC impurities such as organic solvents and are readily dispersed in a wide
CC range of dispersing agents
XX
SQ Sequence 15 AA;
Query Match 73.2%; Score 71; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 LNSKIAFKIVSQEPA 19
Db 1 LNSKIAFKIVSQEPA 15
RESULT 6
AAO17033
ID AAO17033 standard; peptide; 15 AA.
XX
AC AAO17033;
XX
DT 29-MAY-2002 (first entry)
XX
DE Desmoglein 3 residues 190-204 SEQ ID NO: 30.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antinflammatory; antirheumatic; antiarthritis;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX

PR 23-JAN-2003; 2003GB-00001526.
PR 23-JAN-2003; 2003GB-00001527.
PR 23-JAN-2003; 2003GB-00001529.
PR 22-MAR-2003; 2003GB-00006621.
PR 04-APR-2003; 2003WO-GB0001525.
PR 24-MAY-2003; 2003GB-00012062.
PR 01-AUG-2003; 2003WO-GB003285.
PR 03-OCT-2003; 2003GB-00023130.
XX (LORA-) LORANTIS LTD.
PA
XX
XX Bodner MW, Briend ECP, Champion BR, Lennard AC, McKenzie GJ;
PI Tugai T, Ward GA, Young LL;
XX
XX WPI; 2004-534298/51.
XX
XX New product for modulating the immune system, comprises a pharmaceutical
PT support matrix bearing modulators of Notch signaling, and an antigen or
PT antigenic determinant, or a polynucleotide coding for the antigen or
PT determinant.
XX
XX Disclosure; Page 126; 294pp; English.
XX
XX The invention relates to a product comprising (1) a pharmaceutical
CC support matrix for in vivo administration bearing modulators of Notch
CC signalling (especially a Delta or Serrate/Jagged protein or fragment or
CC homologue thereof); and (2) an antigen or antigenic determinant, or
CC polynucleotide encoding an antigen or antigenic determinant. The product
CC acts as a combined preparation for modulation of the immune system or for
CC modulation of an immune response to the antigen or antigenic determinant.
CC The invention also relates to a pharmaceutical composition comprising the
CC product; a particle with a maximum linear dimension of less than 500
CC (preferably 30-70) nm having several bound modulators of Notch signalling
CC; a method of modulating Notch signalling; methods of treating an immune
CC disorder, for reducing an immune response, for promoting immune tolerance
CC in a mammal; and a method for increasing an immune response to a tumour
CC or pathogen antigen or its antigenic determinant in a mammal. The
CC composition and methods are useful for modulating peripheral T-cell
CC activation, for generating regulatory T-cells (T reg), for reducing an
CC immune response to an antigen or antigenic determinant, for promoting
CC immune tolerance to an antigen or antigenic determinant, or for treating
CC tumours, autoimmune disease, allergies or transplant rejection. The
CC present sequence represents an antigenic determinant derived from a human
CC autoantigen which may be used in a product of the invention.
XX
XX Sequence 15 AA;
SQ
Query Match 73.2%; Score 71; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 LNSKIAFKIVSOEPA 19
Db 1 LNSKIAFKIVSOEPA 15
|||||
RESULT 9
ADRA41700
ID ADRA41700 standard; peptide; 15-AA.
XX
AC ADRA41700;
XX
XX 21-OCT-2004 (first entry)
XX
XX Desmoglein 3 (DSG3) (aa 190-204), class II MHC associated autoantigen.
XX DSG3; Desmoglein 3; autoantigen; notch signalling pathway;
KW autoimmune disorder; bystander effect; suppression; DSL domain;
KW EGP domain; Goodpasture's disease; Wegener's granulomatosis; anaemia;
KW thrombocytopenia; gastritis; hepatitis; vasculitis; scleroderma;
KW myositis; arthritis; systemic lupus erythematosus; SLS;
KW Sjogren's syndrome; hepatic fibrosis; liver cirrhosis; thyroiditis;
KW dermatitis; placental dysfunction; eclampsia;

KW inflammatory related gynaecological disease; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; Huntington disease;
KW encephalitis; psychiatric disorder; Down's syndrome; stroke; exogenous;
KW bystander antigen; multiple sclerosis; delta serrate lag;
KW inflammatory bowel disease; notch receptor.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO2004064863-A1.
PN
XX
XX 05-AUG-2004.
PD
XX
XX 23-JAN-2004; 2004WO-GB000263.
PF
XX
XX 23-JAN-2003; 2003GB-00001510.
PR 23-JAN-2003; 2003GB-00001512.
PR 23-JAN-2003; 2003GB-00001513.
PR 23-JAN-2003; 2003GB-00001515.
PR 23-JAN-2003; 2003GB-00001518.
PR 23-JAN-2003; 2003GB-00001519.
PR 23-JAN-2003; 2003GB-00001521.
PR 23-JAN-2003; 2003GB-00001522.
PR 23-JAN-2003; 2003GB-00001524.
PR 23-JAN-2003; 2003GB-00001526.
PR 23-JAN-2003; 2003GB-00001527.
PR 23-JAN-2003; 2003GB-00001529.
PR 04-APR-2003; 2003WO-GB0001525.
PR 24-MAY-2003; 2003GB-00012062.
PR 01-AUG-2003; 2003WO-GB003285.
PR 03-OCT-2003; 2003GB-00023130.
PR 07-JAN-2004; 2004WO-GB000046.
XX (LORA-) LORANTIS LTD.
XX
XX Champion BR, Ragno S, Young LL;
PI WPI; 2004-562091/54.
XX
XX New product having a modulator of the Notch signaling pathway, useful for
PT modulating an immune response in autoimmune disorders, such as anemia,
PT gastritis, hepatitis, scleroderma and myositis.
XX
XX Disclosure; Page 66; 244pp; English.
XX
XX The invention relates to the modulation of immune function through a
CC notch signalling pathway for the prevention of autoimmune diseases. It
CC has been found that the notch signalling pathway provides a bystander
CC effect or bystander suppression effect, which can be used in a wide
CC variety of ways to suppress unwanted immune responses in immune diseases
CC and disorders. Autoimmune diseases are characterised by immune responses
CC that are directed against self antigens. T lymphocytes are activated upon
CC recognition of a self antigen and/or a foreign antigen as a complex with
CC self major histocompatibility complex (MHC) gene products on the surface
CC of antigen presenting cells (APC). The invention provides the method of
CC modulating of an immune response, modulator information and a
CC pharmaceutical kit for suppression of an immune response. The modulator
CC of the notch signalling pathway is an agent which activates the notch
CC receptor or a polynucleotide which codes for such an agent. It comprises
CC a protein or polypeptide comprising a notch ligand DSL (delta serrate
CC lag) domain, notch ligand EGF domain, optionally all or part of a notch
CC ligand N terminal domain, and optionally one or more heterologous amino
CC acid or a polynucleotide sequences. The modulator can be a fusion protein
CC comprising a segment of a notch ligand extracellular domain and an
CC immunoglobulin Fc segment. The disorders include Goodpasture's disease,
CC Wegener's granulomatosis, autoimmune anaemia, thrombocytopenia,
CC gastritis, autoimmune hepatitis, inflammatory bowel disease, autoimmune
CC vasculitis, scleroderma, myositis, autoimmune arthritis, Systemic Lupus
CC Erythematosus (SLE) or Sjogren's syndrome, hepatic fibrosis, liver
CC cirrhosis, thyroiditis, dermatitis, placental dysfunction, eclampsia,
CC inflammatory related gynaecological diseases, neurodegenerative disorders
CC (such as Alzheimer's disease, Parkinson's disease, Huntington disease)
CC encephalitis, psychiatric disorders, Down's syndrome, stroke, multiple

CC sclerosis, etc. The invention discloses a method for generating immune
 CC suppression at a disease locus by administering an exogenous antigen. It
 CC also provides the use of modulator or activator of notch signaling in
 CC simultaneous, separate or sequential combination with a bystander antigen
 CC or antigenic determinant for reducing an immune response to a target
 CC antigen. The presented protein sequence is the desmoglein 3 (DSG3) (aa
 CC 190-204), class II MHC associated autoantigen.
 XX
 SQ Sequence 15 AA;

Query Match 73.2%; Score 71; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
 |||||
 DB 1 LNSKIAFKIVSQEPA 15

RESULT 10
 ADS14310
 ID ADS14310 standard; peptide; 15 AA.
 XX
 AC ADS14310;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Desmoglein 3 antigenic peptide residues 190-204.

XX Cytostatic; Immunosuppressive; Antidiabetic; Neuroprotective;
 KW Antiarthritic; Antirheumatic; Antiallergic; Vaccine; Notch signaling;
 KW Notch; Notch ligand; Delta protein; Serrate protein; Jagged protein;
 KW multiple sclerosis; rheumatoid arthritis; diabetes; allergy;
 KW immune disorder; autoimmune disease; graft rejection; cancer;
 KW organ transplant; desmoglein 3.

XX Unidentified.

XX WO2004083372-A2.

XX 30-SEP-2004.

XX 22-MAR-2004; 2004WO-GB001229.

XX 21-MAR-2003; 2003GB-00006582.

XX 21-MAR-2003; 2003GB-00006583.

XX 22-MAR-2003; 2003GB-00006621.

XX 22-MAR-2003; 2003GB-00006622.

XX 22-MAR-2003; 2003GB-00006624.

XX 22-MAR-2003; 2003GB-00006626.

XX 22-MAR-2003; 2003GB-00006640.

XX 22-MAR-2003; 2003GB-00006644.

XX 22-MAR-2003; 2003GB-00006650.

XX 22-MAR-2003; 2003GB-00006651.

XX 22-MAR-2003; 2003GB-00006654.

XX (LORA-) LORANTIS LTD.

XX Champion BR, Ragno S;

XX WPI; 2004-709927/69.

XX Particle capable of being inserted into or taken up by cell useful for
 PT modulating immune response to antigen in subject, comprises
 PT polynucleotide coding for modulator of Notch signaling, and
 PT polynucleotide coding for antigen.

XX Disclosure; Page 118; 278pp; English.

XX The present invention relates to a particle (i) capable of being inserted
 CC into or taken up by a cell comprising (i) a polynucleotide coding for a
 CC modulator of Notch signaling, and (ii) a polynucleotide coding for an
 CC antigen or antigenic determinant. The first polynucleotide sequence codes

CC for a Notch ligand such as a Delta or Serrate/Jagged protein or its
 CC fragment, derivative, homologue, analogue or allelic variant, or for a
 CC protein which comprises a Notch ligand DSL domain and at least one Notch
 CC ligand EGF-like domain and optionally a membrane binding or transmembrane
 CC domain. The first and second sequences are operably linked to one or more
 CC promoters or enhancers or polyadenylation sequences. The antigen or
 CC antigenic determinant is an allergen, autoantigen, Major
 CC Histocompatibility complex (MHC) (transplant) antigen, pathogen antigen,
 CC tumour antigen or their antigenic determinant. (i) is useful for
 CC modulating an immune response to an antigen in a subject. Pharmaceutical
 CC compositions comprising (i) are useful for treating or preventing
 CC conditions mediated by T-cells, such as multiple sclerosis, rheumatoid
 CC arthritis, diabetes, allergy, for treating immune disorders such as
 CC autoimmune diseases of graft rejection such as allograft rejection,
 CC treating cancer and organ transplants. The present sequence is a
 CC desmoglein 3 antigenic peptide (a class II MHC-associated autoantigen
 CC peptide), which can be used as an antigen to prepare the particle of the
 CC invention.
 XX
 SQ Sequence 15 AA;

Query Match 73.2%; Score 71; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
 |||||
 DB 1 LNSKIAFKIVSQEPA 15

RESULT 11

ABP57835

ID ABP57835 standard; protein; 15 AA.

XX AC ABP57835;

XX 03-FEB-2003 (first entry)

XX A. marginale Msp-2 fragment.

XX Ehrlichiae; Rickettsiae; pathogen; antimicrobial; vaccine;
 KW immune response; Msp-2.

XX Anaplasma marginale.

XX WO200283724-A1.

XX 24-OCT-2002.

XX 11-APR-2002; 2002WO-AU000468.

XX 12-APR-2001; 2001AU-00004400.

XX 10-SEP-2001; 2001AU-00007597.

XX 01-MAR-2002; 2002AU-00000861.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX Riding GA, Hope MA, Willadsen P;

XX WPI; 2003-093001/08.

XX New vaccine, useful for the manufacture of a medicament for raising an
 PT immune response against Ehrlichiae and/or Rickettsiae pathogens.

XX Example 1; Page 32; 102pp; English.

XX The invention relates to a novel vaccine comprising at least one
 CC polypeptide which raises an immune response against Ehrlichiae and/or
 CC Rickettsiae pathogens when administered to a subject. The vaccine of the
 CC invention has antimicrobial activity. The vaccine is useful for the
 CC manufacture of a medicament for raising an immune response against
 CC Ehrlichiae and/or Rickettsiae pathogens or for treating or preventing
 CC Ehrlichiae or Rickettsiae infection in a subject. The present sequence

CC represents a fragment of the A. marginale Map-2 polypeptide
XX
SQ Sequence 15 AA;

Query Match 33.0%; Score 32; DB 6; Length 15;
Best Local Similarity 40.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 4 HLNKIAFKI 13
| : | | : : :
Db 5 HINPKAYRV 14

RESULT 12

ABB04739
ID ABB04739 standard; peptide; 15 AA.

XX
AC ABB04739;

XX
DT 11-MAR-2002 (first entry)

XX Human heparan sulfate 3-O-sulfotransferase 19 peptide SEQ ID NO:7.

XX Human; heparan sulfate 3-O-sulfotransferase 19; haemorrhagic disease;
KW thrombus embolism; haemopathy; myocardial infarction; tumour;
KW inflammation; immunological disease; HIV infection.

XX Homo sapiens.

XX CN1311305-A.

XX PD 05-SEP-2001.

XX PF 02-MAR-2000; 2000CN-00111794.

XX PR 02-MAR-2000; 2000CN-00111794.

XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX PI Mao Y, Xie Y;

XX DR WPI; 2002-049921/07.

XX New polypeptide-heparan sulfate 3-O-sulfotransferase 19 and
PT polynucleotide for coding such polypeptide.

XX Example 6; Page 19 (Disclosure); 34pp; Chinese.

XX The present invention describes human heparan sulfate 3-O-
CC sulfotransferase 19 protein (I). The present invention also describes a
CC method of applying (I) in the treatment of various diseases, such as
CC haemorrhagic diseases, thrombus embolism, other haemopathy, myocardial
CC infarction, various tumours, inflammation, immunological diseases and HIV
CC infection. The present invention also describes the agonist resisting (I)
CC and its treatment effect. The present sequence represents the N-terminal
CC peptide of (I) which is used in an example from the present invention

XX Sequence 15 AA;

Query Match 32.0%; Score 31; DB 5; Length 15;
Best Local Similarity 42.9%; Pred. No. 2.4e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 5 LNSKIAFKIVSQEP 18
: | | | : : | |
Db 1 MNSSIKLLIIVREP 14

RESULT 13

ADJ34233
ID ADJ34233 standard; protein; 15 AA.

XX
AC ADJ34233;

XX
DT

06-MAY-2004 (first entry)

XX Human secreted protein NOV9e.

XX Human; NOVX; secreted protein; cancer; diabetes; obesity;
KW endocrine disorder; CNS disorder; inflammatory disorder; gene therapy.

XX Homo sapiens.

XX WO2004000997-A2.

XX PD 31-DEC-2003.

XX PF 04-JUN-2003; 2003WO-US017512.

XX PR 19-MAR-2002; 2002US-0365491P.

XX PR 04-JUN-2002; 2002US-0385504P.

XX PR 05-JUN-2002; 2002US-0386041P.

XX PR 06-JUN-2002; 2002US-0386453P.

XX PR 06-JUN-2002; 2002US-0386974P.

XX PR 07-JUN-2002; 2002US-0386816P.

XX PR 07-JUN-2002; 2002US-0387002P.

XX PR 10-JUN-2002; 2002US-0387540P.

XX PR 11-JUN-2002; 2002US-0387659P.

XX PR 12-JUN-2002; 2002US-0387934P.

XX PR 13-JUN-2002; 2002US-0389123P.

XX PR 17-JUN-2002; 2002US-0389729P.

XX PR 17-JUN-2002; 2002US-0389742P.

XX PR 19-JUN-2002; 2002US-0390006P.

XX PR 17-JUL-2002; 2002US-0396706P.

XX PR 12-AUG-2002; 2002US-0402832P.

XX PR 13-AUG-2002; 2002US-0403486P.

XX PR 14-AUG-2002; 2002US-0403522P.

XX PR 15-AUG-2002; 2002US-0403748P.

XX PR 06-NOV-2002; 2002US-0387037P.

XX PR 03-JUN-2003; 2003US-0045424E.

XX PA (CURA-) CURAGEN CORP.

XX Anderson DW, Boldog FL, Burgess CE, Casman SJ, Edinger SR;

XX Eisen A, Ellerman K, Gerlach VL, Gorman L, Guo X, Gusev VY, Ji W;

XX Li L, Macdougall JR, Malyankar UM, Millet I, Ort T, Padigar M;

XX Prayaga SK, Patturajan M, Pena CEA, Peyman JA, Rieger DK;

XX Rothenberg ME, Sciore P, Shenoy SG, Smithson G, Spytak KA, Stone DJ;

XX Taupier RJ, Tchernev VT, Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;

XX WPI; 2004-082483/08.

XX N-PSDB; ADJ34232.

XX New isolated NOVX polypeptides useful for treating, preventing and
PT diagnosing pathological conditions with NOVX-associated disorders, such
PT as cancer, obesity, diabetes and inflammatory or CNS diseases.

XX Claim 1; SEQ ID NO 122; 418pp; English.

XX The invention relates to a new isolated polypeptide (designated NOVX)

CC comprising one of 141 fully defined sequences, their mature forms, a

CC protein comprising one or more conservative substitutions or having at

CC least 95% identity to one of the 141 proteins. Also included are a

CC composition comprising NOVX (or a NOVX nucleic acid molecule (NA)), a kit

CC comprising the composition of NOVX in one or more containers, an isolated

CC nucleic acid molecule encoding a NOVX protein, producing NOVX (comprising

CC culturing a cell under conditions that lead to expression of the

CC polypeptide, where the cell comprises a vector comprising NOVX NA),

CC identifying an agent that binds to NOVX, identifying a potential

CC therapeutic agent for use in the treatment of a pathology that is related

CC to aberrant expression or physiological interactions of NOVX, screening

CC for a modulator of activity of or latency or predisposition to a

CC pathology associated with NOVX, modulating the activity of NOVX, treating

CC or preventing a pathology associated with NOVX, treating a pathological

CC state in a mammal, a vector comprising the NOVX nucleic acid molecule, a

CC cell comprising the vector, an antibody that immunospecifically binds to

CC NOVX, determining the presence or amount of NOVX or the nucleic acid
 CC molecule in a sample, and determining the presence of or predisposition
 CC to a disease associated with altered levels of expression of NOVX or the
 CC nucleic acid molecule in a first mammalian subject. The methods and
 CC compositions of the present invention are useful for the diagnosis and
 CC treatment of disorders associated with aberrant expression or activity of
 CC the NOVX polypeptide, such as cancer, diabetes, obesity, and endocrine,
 CC CNS and inflammatory disorders. They can also be used in various
 CC detection and screening assays, chromosome mapping, tissue typing, gene
 CC therapy and predictive medicine. The present sequence represents a NOVX
 CC protein.

CC Sequence 15 AA;

Query Match 32.0%; Score 31; DB 8; Length 15;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 NSKIAPKIV 14
 |||:|
 Db 2 NSKVAFSAV 10

RESULT 14

ID ADJ34231 standard; protein; 16 AA.

AC ADJ34231;

DT 06-MAY-2004 (first entry)

DE Human secreted protein NOV9d.

DE Human; NOVX; secreted protein; cancer; diabetes; obesity;
 KW endocrine disorder; CNS disorder; inflammatory disorder; gene therapy.

OS Homo sapiens.

PN WO2004000997-A2.

PD 31-DEC-2003.

PF 04-JUN-2003; 2003WO-US017512.

PR 19-MAR-2002; 2002US-0365491P.

PR 04-JUN-2002; 2002US-0385504P.

PR 05-JUN-2002; 2002US-0386041P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386974P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0387002P.

PR 10-JUN-2002; 2002US-0387540P.

PR 11-JUN-2002; 2002US-0387659P.

PR 12-JUN-2002; 2002US-0387934P.

PR 13-JUN-2002; 2002US-0389123P.

PR 17-JUN-2002; 2002US-0389729P.

PR 17-JUN-2002; 2002US-0389742P.

PR 19-JUN-2002; 2002US-0390006P.

PR 17-JUL-2002; 2002US-0396706P.

PR 12-AUG-2002; 2002US-0402832P.

PR 13-AUG-2002; 2002US-0403486P.

PR 14-AUG-2002; 2002US-0403522P.

PI Taupier RJ, Tchernev VT, Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;
 XX WPI; 2004-082483/08.
 DR N-PSDB; ADJ34230.

XX New isolated NOVX polypeptides useful for treating, preventing and
 XX diagnosing pathological conditions with NOVX-associated disorders, such
 XX as cancer, obesity, diabetes and inflammatory or CNS diseases.

PS Claim 1; SEQ ID NO 120; 418pp; English.

XX The invention relates to a new isolated polypeptide (designated NOVX)
 CC comprising one of 141 fully defined sequences, their mature forms, a
 CC protein comprising one or more conservative substitutions or having at
 CC least 95% identity to one of the 141 proteins. Also included are a
 CC composition comprising NOVX (or a NOVX nucleic acid molecule (NA)), a kit
 CC comprising the composition of NOVX in one or more containers, an isolated
 CC nucleic acid molecule encoding a NOVX protein, producing NOVX (comprising
 CC culturing a cell under conditions that lead to expression of the
 CC polypeptide, where the cell comprises a vector comprising NOVX NA),
 CC identifying an agent that binds to NOVX, identifying a potential
 CC therapeutic agent for use in the treatment of a pathology that is related
 CC to aberrant expression or physiological interactions of NOVX, screening
 CC for a modulator of activity of or latency or predisposition to a
 CC pathology associated with NOVX, modulating the activity of NOVX, treating
 CC or preventing a pathology associated with NOVX, treating a pathological
 CC state in a mammal, a vector comprising the NOVX nucleic acid molecule, a
 CC cell comprising the vector, an antibody that immunospecifically binds to
 CC NOVX, determining the presence or amount of NOVX or the nucleic acid
 CC molecule in a sample, and determining the presence of or predisposition
 CC to a disease associated with altered levels of expression of NOVX or the
 CC nucleic acid molecule in a first mammalian subject. The methods and
 CC compositions of the present invention are useful for the diagnosis and
 CC treatment of disorders associated with aberrant expression or activity of
 CC the NOVX polypeptide, such as cancer, diabetes, obesity, and endocrine,
 CC CNS and inflammatory disorders. They can also be used in various
 CC detection and screening assays, chromosome mapping, tissue typing, gene
 CC therapy and predictive medicine. The present sequence represents a NOVX
 CC protein.

XX Sequence 16 AA;

Query Match 32.0%; Score 31; DB 8; Length 16;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 NSKIAPKIV 14
 |||:|
 Db 3 NSKVAFSAV 11

RESULT 15

ID AAR21981 standard; peptide; 17 AA.

AC AAR21981;

DT 25-MAR-2003 (revised)

DT 25-JUN-1992 (first entry)

DE hIL-1 binding peptide.

DE Human Interleukin-1; inflammatory diseases; rheumatoid arthritis;
 KW osteoporosis.

OS Synthetic.

PN EP474141-A.

PD 11-MAR-1992.

PF 30-AUG-1991; 91EP-00114678.

PR 03-SEP-1990; 90JP-00233571.
PR 24-NOV-1990; 90JP-00320000.
PR 26-NOV-1990; 90JP-00324956.
XX
PA (KURS) KURARAY CO LTD.
XX
PI Suda T, Abe E, Tanihara M, Fujiwara C;
XX
DR WPI; 1992-081703/11.
XX
PT New peptide(s) with specific binding activity for human interleukin-1 -
PT for treating inflammatory diseases caused by excess interleukin-1 prodn.
PT e.g. rheumatoid arthritis and osteoporosis.
XX
PS Claim 3; Page 6; 9pp; English.
XX
CC The peptide and its salts has high binding activity for human interleukin
CC -1. The peptides can be made by solid phase synthesis. IL-1 is an
CC important factor in immune response and is involved in inflammatory
CC diseases. The peptide inhibits IL-1 activity without binding to the IL-1
CC receptor and is used in treating diseases arising from IL-1, e.g.
CC rheumatoid arthritis and osteoporosis. See also AAR21978-R21993. (Updated
CC on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 17 AA;

Query Match 32.0%; Score 31; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSOEP 18
| | | | |
Db 4 LRKISAKFVNEP 17

Search completed: March 15, 2005, 12:09:40
Job time : 162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 12:12:41 ; Search time 136 Seconds
(without alignments)
46.081 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 270265

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09E_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	73.2	15	9	US-09-909-460-24
2	71	73.2	15	11	US-09-872-836-24
3	71	73.2	15	17	US-10-758-970-24
4	30.5	31.4	19	14	US-10-153-344-10
5	29	29.9	15	13	US-10-103-395-76
6	29	29.9	15	14	US-10-186-867-30
7	29	29.9	16	14	US-10-225-567A-1162
8	29	29.9	18	14	US-10-029-386-33316
9	28	28.9	18	14	US-10-043-487-406
10	28	28.9	18	14	US-10-244-488-21
11	28	28.9	18	14	US-10-244-488-28
12	28	28.9	19	9	US-09-864-761-44631
13	27	27.8	14	16	US-10-712-425-1268

14	27	27.8	17	9	US-09-946-175-6	Sequence 6, Appli
15	26	26.8	9	14	US-10-357-175-101	Sequence 101, App
16	26	26.8	9	15	US-10-455-720-101	Sequence 101, App
17	26	26.8	9	15	US-10-334-726-187	Sequence 187, App
18	26	26.8	9	15	US-10-024-652-40	Sequence 40, Appl
19	26	26.8	9	15	US-10-024-652-601	Sequence 601, App
20	26	26.8	9	15	US-10-024-652-894	Sequence 894, App
21	26	26.8	9	15	US-10-024-652-983	Sequence 983, App
22	26	26.8	9	15	US-10-024-652-1068	Sequence 1068, Ap
23	26	26.8	9	15	US-10-024-652-1391	Sequence 1391, Ap
24	26	26.8	9	15	US-10-024-652-1427	Sequence 1427, Ap
25	26	26.8	9	15	US-10-024-652-1470	Sequence 1470, Ap
26	26	26.8	10	15	US-10-024-652-63	Sequence 63, Appl
27	26	26.8	10	15	US-10-024-652-65	Sequence 65, Appl
28	26	26.8	10	15	US-10-024-652-200	Sequence 200, App
29	26	26.8	10	15	US-10-024-652-274	Sequence 274, App
30	26	26.8	10	15	US-10-024-652-372	Sequence 372, App
31	26	26.8	10	15	US-10-024-652-569	Sequence 569, App
32	26	26.8	10	15	US-10-024-652-660	Sequence 660, App
33	26	26.8	10	15	US-10-024-652-686	Sequence 686, App
34	26	26.8	10	15	US-10-024-652-1592	Sequence 1592, Ap
35	26	26.8	10	15	US-10-024-652-1618	Sequence 1618, Ap
36	26	26.8	10	15	US-10-024-652-1619	Sequence 1619, Ap
37	26	26.8	10	15	US-10-024-652-1701	Sequence 1701, Ap
38	26	26.8	10	15	US-10-024-652-1709	Sequence 1709, Ap
39	26	26.8	10	15	US-10-024-652-1730	Sequence 1730, Ap
40	26	26.8	10	15	US-10-024-652-1818	Sequence 1818, Ap
41	26	26.8	10	15	US-10-024-652-1896	Sequence 1896, Ap
42	26	26.8	10	15	US-10-024-652-1904	Sequence 1904, Ap
43	26	26.8	11	9	US-09-795-006A-137	Sequence 137, App
44	26	26.8	12	14	US-10-286-457-34	Sequence 34, Appl
45	26	26.8	13	14	US-10-403-980-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-909-460-24
; Sequence 24, Application US/09909460
; Publication No. US20020182258A1
; GENERAL INFORMATION:
; APPLICANT: Lunsford, Lynn B.
; APPLICANT: Putnam, David
; APPLICANT: Hedley, Mary Lynn
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF NUCLEIC
; TITLE OF INVENTION: ACID
; FILE REFERENCE: 08191/014001
; CURRENT APPLICATION NUMBER: US/09/909,460
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-05-27
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-460-24

Query Match 73.2%; Score 71; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
|||||
DB 1 LNSKIAFKIVSQEPA 15

RESULT 2
US-09-872-836-24
; Sequence 24, Application US/09872836
; Publication No. US20040142475A1

GENERAL INFORMATION:
APPLICANT: Barmen, Shikha P.
APPLICANT: McKeever, Una
APPLICANT: Hedley, Mary Lynne
TITLE OF INVENTION: DELIVERY SYSTEMS FOR BIOACTIVE AGENTS
FILE REFERENCE: 08191-018001
CURRENT APPLICATION NUMBER: US/09/872,836
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: US 60/208,830
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 120
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
US-09-872-836-24

Query Match 73.2%; Score 71; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.5e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 15; Conservative 0;

Qy 5 LNSKIAFKIVSQEPA 19
Db 1 LNSKIAFKIVSQEPA 15

RESULT 3
US-10-758-970-24
Sequence 24, Application US/10758970
Publication No. US20050037086A1
GENERAL INFORMATION:
APPLICANT: Hedley, Mary Lynne
APPLICANT: Hsu, Yung-Yueh
APPLICANT: Tyo, Michael
TITLE OF INVENTION: CONTINUOUS-FLOW METHOD FOR PREPARING MICROPARTICLES
FILE REFERENCE: 08191-012001
CURRENT APPLICATION NUMBER: US/10/758,970
CURRENT FILING DATE: 2004-01-16
PRIOR APPLICATION NUMBER: US/09/715,708A
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: US 60/166,516
PRIOR FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 109
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
US-10-758-970-24

Query Match 73.2%; Score 71; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.5e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 15; Conservative 0;

Qy 5 LNSKIAFKIVSQEPA 19
Db 1 LNSKIAFKIVSQEPA 15

RESULT 4
US-10-153-344-10
Sequence 10, Application US/10153344
Publication No. US20030004124A1
GENERAL INFORMATION:
APPLICANT: ROTHMAN, JOEL
APPLICANT: BLOSS, TIM
APPLICANT: WITZE, ERIC
TITLE OF INVENTION: BTF3: AN INHIBITOR OF APOPTOSIS
FILE REFERENCE: 407T-300410US
CURRENT APPLICATION NUMBER: US/10/153,344
CURRENT FILING DATE: 2002-08-27
PRIOR APPLICATION NUMBER: US 60/292,559

PRIOR FILING DATE: 2001-05-21
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 19
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-153-344-10

Query Match 31.4%; Score 30.5; DB 14; Length 19;
Best Local Similarity 47.1%; Pred. No. 4.3e+02; Mismatches 8; Conservative 4; Indels 3; Gaps 1;
Matches 8; Conservative 4;

Qy 3 NMLNSKIAFKIVSQEPA 19
Db 2 DHLRAK---KILSREDA 15

RESULT 5
US-10-103-395-76
Sequence 76, Application US/10103395
Publication No. US20020160019A1
GENERAL INFORMATION:
APPLICANT: EPIMUNE, Inc.
APPLICANT: Sette, Alessandro
APPLICANT: Sidney, John
APPLICANT: Southwood, Scott
TITLE OF INVENTION: IDENTIFICATION OF BROADLY REACTIVE DR

TITLE OF INVENTION: RESTRICTED EPITOPES
FILE REFERENCE: 39963-20016.01
CURRENT APPLICATION NUMBER: US/10/103,395
CURRENT FILING DATE: 2003-01-03
PRIOR APPLICATION NUMBER: US 09/009,953
PRIOR FILING DATE: 1998-01-21
PRIOR APPLICATION NUMBER: PCT/US98/01373
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: US 60/036,713
PRIOR FILING DATE: 1997-01-23
PRIOR APPLICATION NUMBER: US 60/037,432
PRIOR FILING DATE: 1997-02-07
NUMBER OF SEQ ID NOS: 274
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 76
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
US-10-103-395-76

Query Match 29.9%; Score 29; DB 13; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.9e+02; Mismatches 6; Conservative 1; Indels 0; Gaps 0;
Matches 6; Conservative 1;

Qy 11 FKIVSQEPA 19
Db 4 FKIGSDPA 12

RESULT 6
US-10-186-867-30
Sequence 30, Application US/10186867
Publication No. US20030100508A1
GENERAL INFORMATION:
APPLICANT: Simon-Haldi, Maryline
APPLICANT: Schachner, Melitta
APPLICANT: Neuberger, Timothy
APPLICANT: Uri, Herzberg
TITLE OF INVENTION: CARBOHYDRATE EPITOPE MIMIC COMPOUNDS AND USES THEREOF
FILE REFERENCE: 1094-1-008NCIP
CURRENT APPLICATION NUMBER: US/10/186,867
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: US/09/511,956
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/121,327

```
; PRIOR FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: 60/155,492
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 30
; LENGTH: 15
; TYPE: PRT
; ORGANISM: m13 library
US-10-186-867-30

Query Match      29.9%; Score 29; DB 14; Length 15;
Best Local Similarity 45.5%; Pred. No. 5.9e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      5 LNSKIAFKIVS 15
      |::: |::|
Db      2 LNRLLFRIVS 12
      |::: |::|

RESULT 7
US-10-225-567A-1162
; Sequence 1162, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1162
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-1162

Query Match      29.9%; Score 29; DB 14; Length 16;
Best Local Similarity 60.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      4 HLNSKIAFKI 13
      || |::| :|
Db      5 HLESKISKRI 14
      || |::| :|

RESULT 8
US-10-029-386-33316
; Sequence 33316, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33316
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000051.2
```

```
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.44
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.43
US-10-029-386-33316

Query Match      29.9%; Score 29; DB 14; Length 18;
Best Local Similarity 83.3%; Pred. No. 7.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4 HLNSKI 9
      ||::|:
Db      9 HLNSKL 14
      ||::|:

RESULT 9
US-10-043-487-406
; Sequence 406, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptid
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 406
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Shigella Flexneri
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: MISC_FEATURE
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(8)
; OTHER INFORMATION: MISC_FEATURE
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13)..(13)
; OTHER INFORMATION: MISC_FEATURE
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15)..(15)
; OTHER INFORMATION: MISC_FEATURE
US-10-043-487-406

Query Match      28.9%; Score 28; DB 14; Length 18;
Best Local Similarity 35.7%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      5 LNSKIAFKIVSQEP 18
      |::|:|:|:|
Db      5 LRXXVAIEKLXQEP 18
      |::|:|:|:|

RESULT 10
US-10-244-488-21
; Sequence 21, Application US/10244488
; Publication No. US20030104495A1
; GENERAL INFORMATION:
; APPLICANT: Carr, Daniel W
; APPLICANT: Vijayaraghavan, Srinivasan
; TITLE OF INVENTION: Sperm-specific AKAP Protein Genes and Uses
; FILE REFERENCE: 98700
; CURRENT APPLICATION NUMBER: US/10/244,488
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US/09/268,480
; PRIOR FILING DATE: 1999-03-16
```

NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 21
 LENGTH: 18
 TYPE: PRT
 ORGANISM: Bos sp.
 US-10-244-488-21

Query Match 28.9%; Score 28; DB 14; Length 18;
 Best Local Similarity 55.6%; Pred. No. 1.1e+03;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 8 KIAFKIVSQ 16
 :||:|
 Db 4 RAAFLIISQ 12
 :||:|

RESULT 11
 US-10-244-488-28
 Sequence 28, Application US/10244488
 Publication No. US2003010495A1
 GENERAL INFORMATION:
 APPLICANT: Carr, Daniel W
 TITLE OF INVENTION: Sperm-specific AKAP Protein Genes and Uses
 FILE REFERENCE: 98700
 CURRENT APPLICATION NUMBER: US/10/244,488
 CURRENT FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: US/09/268,480
 PRIOR FILING DATE: 1999-03-16
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 28
 LENGTH: 18
 TYPE: PRT
 ORGANISM: Bos sp.
 US-10-244-488-28

Query Match 28.9%; Score 28; DB 14; Length 18;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQ 16
 ||:|:
 Db 1 LEEKVAALVSQ 12
 ||:|:

RESULT 12
 US-09-864-761-44631
 Sequence 44631, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aecomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/006666
 PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/006670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 44631
 LENGTH: 19
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC018528.3
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.75
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.84
 US-09-864-761-44631

Query Match 28.9%; Score 28; DB 9; Length 19;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 NSKIAFKIVSQE 17
 ||:|:
 Db 6 NYKLSVKFSQOE 17
 ||:|:

RESULT 13
 US-10-712-425-1268
 Sequence 1268, Application US/10712425
 Publication No. US20040180380A1
 GENERAL INFORMATION:
 APPLICANT: LEE, FRANK D.
 APPLICANT: MENG, XUN
 APPLICANT: LIVINGSTON, DAVID
 TITLE OF INVENTION: PROTEOME EPITOPE TAGS AND METHODS OF USE THEREOF IN PROTEIN
 FILE REFERENCE: ENGE-P02-001
 CURRENT APPLICATION NUMBER: US/10/712,425
 CURRENT FILING DATE: 2003-11-13
 PRIOR APPLICATION NUMBER: 60/379,626
 PRIOR FILING DATE: 2002-05-10
 PRIOR APPLICATION NUMBER: 60/393,137
 PRIOR FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: 60/393,197
 PRIOR FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: 60/393,211
 PRIOR FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: 60/393,223
 PRIOR FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: 60/393,233

```
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,235
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,280
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/430,948
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/433,319
; PRIOR FILING DATE: 2002-12-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1268
; TYPE: PRT
; LENGTH: 14
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Illustrative peptide having high affinity for its antigen
US-10-712-425-1268
```

```
Query Match      27.8%; Score 27; DB 16; Length 14;
Best Local Similarity 55.6%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 PNHLSKIA 10
      |||||
Db      4 PEDLNOKLA 12
```

```
RESULT 14
US-09-946-175-6
; Sequence 6, Application US/09946175
; Patent No. US20020106671A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert
; TITLE OF INVENTION: IDENTIFICATION OF A CAPACITATIVE CALCIUM
; CHANNEL IN ANTIGEN PRESENTING CELLS AND USES THEREOF
; FILE REFERENCE: Mlsn1(310800)
; CURRENT APPLICATION NUMBER: US/09/946,175
; CURRENT FILING DATE: 2001-09-05
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-175-6
```

```
Query Match      27.8%; Score 27; DB 9; Length 17;
Best Local Similarity 47.1%; Pred. No. 1.5e+03;
Matches 8; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY      1 EPNHLSKIAPKIVSQE 17
      |||||
Db      3 EPGKLSQKI--KWLQOE 17
```

```
RESULT 15
US-10-357-175-101
; Sequence 101, Application US/10357175
; Publication No. US20030170707A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
; FILE REFERENCE: in Ovarian Carcinoma and Uses Thereof
; FILE REFERENCE: D6192CIP/D/CIP
; CURRENT APPLICATION NUMBER: US/10/357,175
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/650,371
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 101
; LENGTH: 9
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: 283....291
; OTHER INFORMATION: TADG-12 peptide
US-10-357-175-101

Query Match      26.8%; Score 26; DB 14; Length 9;
Best Local Similarity 62.5%; Pred. No. 1.3e+06;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 PNHLSKI 9
      |||||
Db      2 PSHLVEKI 9

Search completed: March 15, 2005, 12:25:15
Job time : 137 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2005, 12:04:30 ; Search time 42 Seconds
(without alignments)
33.770 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLNSKIAPKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 187692

Minimum DB seq length: 0
Maximum DB seq length: 19

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	73.2	15	1	US-08-787-547-24
2	71	73.2	15	2	US-08-400-796-3
3	31	32.0	12	2	US-08-637-759B-448
4	31	32.0	12	3	US-08-871-355A-448
5	31	32.0	12	3	US-09-201-945-448
6	31	32.0	17	1	US-08-239-854-6
7	30	30.9	17	3	US-08-160-604-62
8	29	29.9	15	4	US-09-009-953-76
9	28	28.9	18	4	US-09-268-480-21
10	28	28.9	18	4	US-09-268-480-28
11	27	27.8	8	3	US-08-160-604-65
12	27	27.8	8	3	US-08-160-604-66
13	27	27.8	8	3	US-08-160-604-67
14	27	27.8	19	4	US-09-442-989-30
15	26	26.8	19	3	US-09-518-046-101
16	26	26.8	15	1	US-08-204-656B-16
17	26	26.8	15	1	US-08-470-702-16
18	26	26.8	15	1	US-08-467-831-16
19	26	26.8	17	4	US-10-204-664A-3
20	26	26.8	19	3	US-09-308-935-15
21	25	25.8	10	4	US-09-490-702B-57
22	25	25.8	12	4	US-08-634-332A-21
23	25	25.8	12	4	US-08-634-332A-22
24	25	25.8	13	1	US-08-548-540-152
25	25	25.8	13	1	US-08-191-338A-29
26	25	25.8	13	1	US-08-191-338A-30
27	25	25.8	13	1	US-08-191-338A-31

28	25	25.8	13	5	PCT-US96-09809-152	Sequence 152, App
29	25	25.8	15	1	US-08-049-783-11	Sequence 11, Appl
30	25	25.8	15	1	US-08-158-232-22	Sequence 22, Appl
31	25	25.8	15	1	US-08-304-626-22	Sequence 22, Appl
32	25	25.8	15	1	US-08-316-301A-23	Sequence 23, Appl
33	25	25.8	15	2	US-08-611-928-22	Sequence 22, Appl
34	25	25.8	15	3	US-09-224-024-5	Sequence 5, Appl
35	25	25.8	15	3	US-09-173-891-22	Sequence 22, Appl
36	25	25.8	15	3	US-09-076-137-23	Sequence 23, Appl
37	25	25.8	15	4	US-09-009-953-23	Sequence 23, Appl
38	25	25.8	15	4	US-09-009-953-46	Sequence 46, Appl
39	25	25.8	15	4	US-09-009-953-134	Sequence 134, App
40	25	25.8	15	4	US-09-009-953-223	Sequence 223, App
41	25	25.8	15	4	US-09-009-953-224	Sequence 224, App
42	25	25.8	15	4	US-09-311-784A-418	Sequence 418, App
43	25	25.8	15	4	US-09-738-363-23	Sequence 23, Appl
44	25	25.8	15	5	PCT-US92-03624-23	Sequence 23, Appl
45	25	25.8	15	5	PCT-US94-07902-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-787-547-24
; Sequence 24, Application US/08787547
; Patent No. 5783567
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; APPLICANT: Curley, Joanne M.
; APPLICANT: Langer, Robert S.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY
; TITLE OF INVENTION: OF NUCLEIC ACID
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,547
; FILING DATE: 22-JAN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 08191/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-787-547-24

Query Match 73.2%; Score 71; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
|||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 2

US-08-400-796-3
; Sequence 3, Application US/08400796
; Patent No. 5874531
; GENERAL INFORMATION:
; APPLICANT: STROMINGER, JACK L.
; APPLICANT: WUCHERPFENNIG, KAI
; TITLE OF INVENTION: IDENTIFICATION OF SELF AND NON-SELF
; TITLE OF INVENTION: ANTIGENS IMPLICATED IN AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,796
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: TWOMEY, MICHAEL J.
; REGISTRATION NUMBER: 38,349
; REFERENCE/DOCKET NUMBER: H0498/7015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS

US-08-400-796-3

Query Match 73.2%; Score 71; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
|||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 3

US-08-637-759B-448
; Sequence 448, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia

; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 448:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-637-759B-448

Query Match 32.0%; Score 31; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 HLNSKIAFKIVS 15
:|:|:|:|:
Db 1 YANNIIAQVVS 12

RESULT 4

US-08-871-355A-448
; Sequence 448, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: RPMs 101 CON
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 448:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-448

Query Match 32.0%; Score 31; DB 3; Length 12;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 HNSKIAPKIVS 15
DB 1 YANNIAFQVVS 12

RESULT 5
US-09-201-945-448
Sequence 448, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMs 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 448:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-448

Query Match 32.0%; Score 31; DB 3; Length 12;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 HNSKIAPKIVS 15
DB 1 YANNIAFQVVS 12

RESULT 6
US-08-239-854-6
Sequence 6, Application US/08239854
Patent No. 5624900
GENERAL INFORMATION:
APPLICANT: Suda, Tatsuo
APPLICANT: Abe, Etsuko
APPLICANT: Tanihara, Masao
APPLICANT: Fujiwara, Chie
TITLE OF INVENTION: Peptide or Its Salts
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/239,854
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/752,426
FILING DATE:
APPLICATION NUMBER: JP 233,571/90
FILING DATE: 03-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 320,000/90
FILING DATE: 24-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 324,956/90
FILING DATE: 26-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5624900man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 363-279-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 521-4500
TELEFAX: (703) 486-2347
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-239-854-6

Query Match 32.0%; Score 31; DB 1; Length 17;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSOEP 18
DB 4 LRIKISAFVNEP 17

RESULT 7
US-08-160-604-62
Sequence 62, Application US/08160604
Patent No. 6232522
GENERAL INFORMATION:

APPLICANT: Harley, John
APPLICANT: James, Judith A.
APPLICANT: Scofield, R. H.
TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,604
FILING DATE: 30-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/867,819
FILING DATE: 13-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/648,205
FILING DATE: 31-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,947
FILING DATE: 31-JAN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRP114CIP(3)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-160-604-62

Query Match 30.9%; Score 30; DB 3; Length 17;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 NHLNSKIAPK 12
Db 5 NHLKSKEVWK 14

RESULT 8
US-09-009-953-76
Sequence 76, Application US/09009953
Patent No. 6413517
GENERAL INFORMATION:
APPLICANT: Sette, Alessandro
TITLE OF INVENTION: Identification of Broadly
REACTIVE DR Restricted Epitopes
NUMBER OF SEQUENCES: 274
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,953
FILING DATE: 21-Jan-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/036,713
FILING DATE: 23-JAN-1997
APPLICATION NUMBER: US 60/037,432
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-011520US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 76:
US-09-009-953-76

Query Match 29.9%; Score 29; DB 4; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 11 FKIVSQEPA 19
Db 4 FKIGSDPA 12

RESULT 9
US-09-268-480-21
Sequence 21, Application US/09268480A
Patent No. 6451528
GENERAL INFORMATION:
APPLICANT: Carr, Daniel W
APPLICANT: Vijayaraghavan, Srinivasan
TITLE OF INVENTION: Sperm-specific AKAP Protein Genes and Uses
FILE REFERENCE: 98700
CURRENT APPLICATION NUMBER: US/09/268,480A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 21
LENGTH: 18
TYPE: PRT
ORGANISM: Bos sp.
US-09-268-480-21

Query Match 28.9%; Score 28; DB 4; Length 18;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 8 KIAFKIVSQ 16
Db 4 RAAFIISQ 12

RESULT 10
US-09-268-480-28
Sequence 28, Application US/09268480A

; Patent No. 6451528
; GENERAL INFORMATION:
; APPLICANT: Carr, Daniel W
; APPLICANT: Vijayaraghavan, Srinivasan
; TITLE OF INVENTION: Sperm-specific AKAP Protein Genes and Uses.
; FILE REFERENCE: 98700
; CURRENT APPLICATION NUMBER: US/09/268,480A
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bos sp.
US-09-268-480-28

Query Match 28.9%; Score 28; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQ 16
| | | | |
DB 1 LEEKVAALVSQ 12

RESULT 11
US-08-160-604-65
; Sequence 65, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,604
; FILING DATE: 30-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/867,819
; FILING DATE: 13-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,205
; FILING DATE: 31-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,947
; FILING DATE: 31-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

Query Match 27.8%; Score 27; DB 3; Length 8;
Best Local Similarity 83.3%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-160-604-65

Query Match 27.8%; Score 27; DB 3; Length 8;
Best Local Similarity 83.3%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NHLNSK 8
| | | | |
DB 3 NHLKSK 8

RESULT 12
US-08-160-604-66
; Sequence 66, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,604
; FILING DATE: 30-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/867,819
; FILING DATE: 13-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,205
; FILING DATE: 31-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,947
; FILING DATE: 31-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-160-604-66

QY 3 NHLNSK 8
 Db 2 NHLKSK 7

RESULT 13
 US-08-160-604-67
 ; Sequence 67, Application US/08160604
 ; Patent No. 6232522
 ; GENERAL INFORMATION:
 ; APPLICANT: Harley, John
 ; APPLICANT: James, Judith A.
 ; APPLICANT: Scofield, R. H.
 ; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
 ; NUMBER OF SEQUENCES: 127
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Patrea L. Pabst
 ; STREET: 1100 Peachtree Street, Suite 2800
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: USA
 ; ZIP: 30309-4530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/160,604
 ; FILING DATE: 30-NOV-1993
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/867,819
 ; FILING DATE: 13-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/648,205
 ; FILING DATE: 31-JAN-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/472,947
 ; FILING DATE: 31-JAN-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pabst, Patrea L.
 ; REGISTRATION NUMBER: 31,284
 ; REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (404)-815-6508
 ; TELEFAX: (404)-815-6555
 ; INFORMATION FOR SEQ ID NO: 67:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; US-08-160-604-67

Query Match 27.8%; Score 27; DB 3; Length 8;
 Best Local Similarity 83.3%; Pred. NO. 4.1e+05;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NHLNSK 8
 Db 1 NHLKSK 6

RESULT 14
 US-09-442-989-30
 ; Sequence 30, Application US/09442989
 ; Patent No. 6569993
 ; GENERAL INFORMATION:

; APPLICANT: Sledeski, Adam W.
 ; APPLICANT: Mencil, James J.
 ; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF RESIN-BOUND CYCLIC
 ; FILE REFERENCE: A3113B-US
 ; CURRENT APPLICATION NUMBER: US/09/442,989
 ; CURRENT FILING DATE: 1999-11-18
 ; EARLIER APPLICATION NUMBER: 60/081,897
 ; EARLIER FILING DATE: 1998-04-15
 ; NUMBER OF SEQ ID NOS: 46
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 30
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (1)
 ; OTHER INFORMATION: FMOC-Lys (BOC)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (2)
 ; OTHER INFORMATION: His (Trt)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (4)
 ; OTHER INFORMATION: Asn (Trt)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (5)
 ; OTHER INFORMATION: Ser (OtBu)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (6)..
 ; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at
 ; OTHER INFORMATION: position 6 and Asp at position 10 are linked by an
 ; OTHER INFORMATION: amide bond."
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (7)
 ; OTHER INFORMATION: Glu (OtBu)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (8)
 ; OTHER INFORMATION: Arg (PMC)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (6)..
 ; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at
 ; OTHER INFORMATION: position 6 and Asp at position 10 are linked by an
 ; OTHER INFORMATION: amide bond."
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (11)
 ; OTHER INFORMATION: Trp (BOC)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (13)
 ; OTHER INFORMATION: Arg (PMC)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (14)
 ; OTHER INFORMATION: Lys (BOC)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (17)
 ; OTHER INFORMATION: Gln (Trt)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (18)
 ; OTHER INFORMATION: Asp (OtBu)
 ; FEATURE:
 ; NAME/KEY: PEPTIDE

; LOCATION: (6)...(10)
; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at
; position 6 and Asp at position 10 are linked by an
; amide bond."
US-09-442-989-30

Query Match 27.8%; Score 27; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 HLNSK 8
| | | | |
Db 2 HLNSK 6

RESULT 15
US-09-518-046-101
; Sequence 101, Application US/09518046
; Patent No. 6294663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof
; FILE REFERENCE: D6192CIP
; CURRENT APPLICATION NUMBER: US/09/518,046
; CURRENT FILING DATE: 2000-03-02
; EARLIER APPLICATION NUMBER: 09/261,416
; EARLIER FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 101
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: 283...291
; OTHER INFORMATION: TADG-12 peptide
US-09-518-046-101

Query Match 26.8%; Score 26; DB 3; Length 9;
Best Local Similarity 62.5%; Pred. No. 4.1e+05;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PNLNSKI 9
| : | | |
Db 2 PSHLVEKI 9

Search completed: March 15, 2005, 12:14:07
Job time : 43 secs

This Page Blank (uspto)